Figure 1

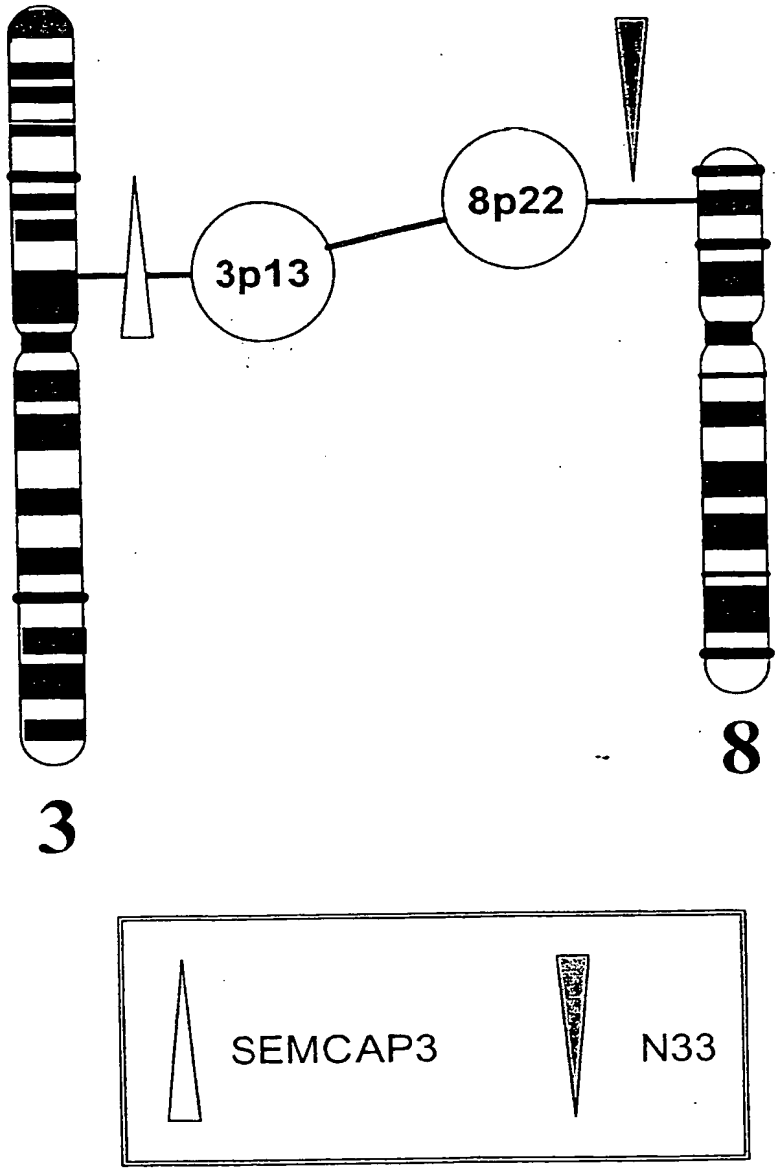


Figure 2

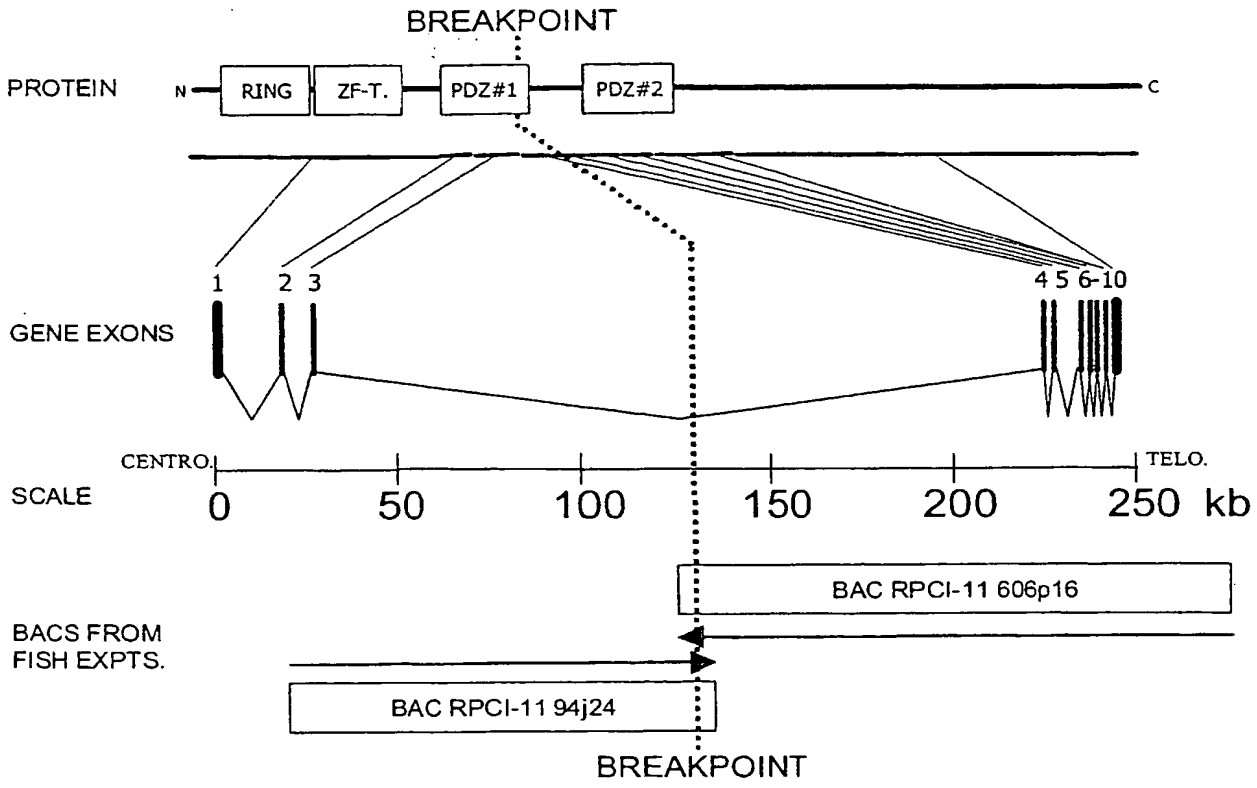

BREAKPOINT
PROTEIN
N
RING
ZF-T.
PDZ#1
PDZ#2
C
1
2 3
4 5 6-10
GENE EXONS
CENTRO.
TELO.
SCALE
0
50
100
150
200
250 kb
BAC RPCI-11 606p16
BACS FROM
FISH EXPTS.
BAC RPCI-11 94j24
BREAKPOINT

Figure 3

```
   1 AAAACTTCCC CGGGTAGATT CACCCACCGG TCCTGGAAAC CTGCTAAATC CTGAAGGTTC
  61 ACAGAACCTC TGGTCAGAAC TGAAGTTGCA GCCGGAGCTT CCCGCAGGCT CTGTAACTTT
 121 CCCTGGAATG AAATAAATAA ATAAAGACCG TAAGTGCTGA GATAGCGGGC CCCAAGATAT
 181 TTTTAGTCCT CTGCAATCAG CCACTAGAGG AAGGGGGAGG GAGAAGGGAG TAAAAAAGTT
 241 TTGATCCGTT CGGGAAGGGG CTCGAAGAGA ACCCTTGGGA GAAAGCAGTA GCCTCAGCTC
 301 CAAACTCAGC GAGCTTTTCT CGGCTGGCGT TTTGTCTCCT ATAGCGTAGA CTGTAAGAGA
 361 ACAGAAAGGA GTTTCCCGAG AAGATTCAGG CTGGCGTCCT GGGCTGGCCC GTCCCTTCTG
 421 GCGAGCCTCA GTGTCCTCCC ACGCGCTTCT GCCTTCCAGC CTCCTCCCTT TTTCGGGGGG
 481 CTGGCGGGAG GCATCCAAGG CACGATGTAT GTGCGCTCGC GCTCGCGCAA ATACGGCCGG
 541 AGGAGTCCTG TTCCTCGGGC ATTTTCCGAG GAAGTCTGGA TCAATTAGGC TCAGTCCGGG
 601 GAGAGCCAGC GAGCGCGCGG GCGGCGTAGC CGGCCTGTCT GGGCCGCCTC GTGGGGAGGG
 661 AGGGGGCGCC CGGCCGCCCG GCGGCGACCC CGGGGCCTGG CCGCCACCAT GGGCTTCGAG
 721 CTGGACCGCT TCGACGGCGA CGTGGACCCG GACCTGAAGT GCGCGCTGTG CCACAAGGTC
 781 CTGGAGGACC CGCTGACCAC GCCGTGCGGC CACGTCTTCT GCGCCGGCTG CGTGCTGCCC
 841 TGGGTGGTGC AGGAGGGCAG CTGCCCGGCG CGCTGCCGCG GTCGCCTGTC GGCCAAAGAG
 901 CTCAACCACG TCCTGCCGCT CAAGCGCCTT ATCCTCAAGC TGGACATCAA GTGCGCGTAC
 961 GCGACGCGCG GCTGCGGCCG GGTGGTCAAG CTGCAGCAGC TGCCGGAGCA CCTCGAGCGC
1021 TGCGACTTCG CGCCCGCGCG CTGTCGCCAC GCGGGTTGCG GCCAGGTGCT GCTGCGGCGC
1081 GACGTGGAGG CGCACATGCG CGACGCGTGC GACGCGCGGC CAGTGGGCCG CTGCCAGGAG
1141 GGCTGCGGGC TACCCTTGAC GCACGGCGAG CAGCGCGCGG GCGGCCACTG CTGCGCGCGA
1201 GCGCTGCGGG CGCACAACGG CGCGCTCCAG GCCCGCCTGG GCGCGCTGCA CAAGGCGCTC
1261 AAGAAGGAGG CGCTGCGCGC TGGGAAGCGC GAGAAGTCGC TGGTGGCCCA GCTGGCCGCG
1321 GCGCAGCTTG AGCTGCAGAT GACCGCGCTG CGCTACCAGA AGAAATTCAC CGAATACAGC
1381 GCGCGCCTCG ACTCGCTCAG CCGCTGCGTG GCCGCGCCGC CCGGCGGCAA GGGCGAAGAA
1441 ACCAAAAGTC TGACTCTTGT CCTGCATCGG GACTCCGGCT CCCTGGGATT CAATATTATT
1501 GGTGGCCGGC CGAGTGTGGA TAACCACGAT GGATCATCCA GTGAAGGAAT CTTTGTATCC
1561 AAGATAGTTG ACAGTGGGCC TGCAGCCAAG GAAGGAGGCC TGCAAATTCA TGACAGGATT
1621 ATTGAGGTCA ACGGCAGAGA CTTATCCAGA GCAACTCATG ACCAGGCTGT GGAAGCTTTC
1681 AAGACAGCCA AGGAGCCCAT AGTGGTGCAG GTGTTGAGAA GAACACCAAG GACCAAAATG
1741 TTCACGCCTC CATCAGAGTC TCAGCTGGTG GACACGGGAA CCCAAACCGA CATCACCTTT
1801 GAACATATCA TGGCCCTCAC TAAGATGTCC TCTCCCAGCC CACCCGTGCT GGATCCCTAT
1861 CTCTTGCCAG AGGAGCATCC CTCAGCCCAT GAATACTACG ATCCAAATGA CTACATTGGA
1921 GACATCCATC AGGAGATGGA CAGGGAGGAG CTGGAGCTGG AGGAAGTGGA CCTCTACAGA
1981 ATGAACAGCC AGGACAAGCT GGGCCTCACT GTGTGCTACC GGACGGACGA TGAAGACGAC
```

2041 ATTGGGATTT ATATCAGTGA GATTGACCCT AACAGCATTG CAGCCAAGGA TGGGCGCATC
2101 CGAGAAGGAG ACCGCATTAT CCAGATTAAT GGGATAGAGG TGCAGAACCG TGAAGAGGCT
2161 GTGGCTCTTC TAACCAGTGA AGAAAATAAA AACTTTTCAT TGCTGATTGC AAGGCCTGAA
2221 CTCCAGCTGG ATGAGGGCTG GATGGATGAT GACAGGAACG ACTTTCTGGA TGACCTGCAC
2281 ATGGACATGC TGGAGGAGCA GCACCACCAG GCCATGCAAT TCACAGCTAG CGTGCTGCAG
2341 CAGAAGAAGC ACGACGAAGA CGGTGGGACC ACAGATACAG CCACCATCTT GTCCAACCAG
2401 CACGAGAAGG ACAGCGGTGT GGGGCGGACC GACGAGAGCA CCCGTAATGA CGAGAGCTCG
2461 GAGCAAGAGA ACAATGGCGA CGACGCCACC GCATCCTCCA ACCCGCTGGC GGGGCAGAGG
2521 AAGCTCACCT GCAGCCAGGA CACCTTGGGC AGCGGCGACC TGCCCTTCAG CAACGAGTCT
2581 TTCATTTCGG CCGACTGCAC GGACGCCGAC TACCTGGGGA TCCCGGTGGA CGAGTGCGAG
2641 CGCTTCCGCG AGCTCCTGGA GCTCAAGTGC CAGGTGAAGA GCGCCACCCC TTACGGCCTG
2701 TACTACCCTA GCGGCCCCCT GGACGCCGGC AAGAGTGACC CTGAGAGCGT GGACAAGGAG
2761 CTGGAGCTGC TGAACGAAGA GCTGCGCAGC ATCGAGCTGG AGTGCCTGAG CATCGTGCGC
2821 GCCCACAAGA TGCAGCAGCT CAAGGAGCAG TACCGCGAGT CCTGGATGCT GCACAACAGC
2881 GGCTTCCGCA ACTACAACAC CAGCATCGAC GTGCGCAGAC ACGAGCTCTC AGATATCACC
2941 GAGCTCCCGG AGAAATCCGA CAAGGACAGC TCGAGCGCCT ACAACACAGG CGAGAGCTGC
3001 CGCAGCACCC CGCTCACCCT GGAGATCTCC CCCGACAACT CCTTGAGGAG AGCGGCGGAG
3061 GGCATCAGCT GCCCGAGCAG CGAAGGGGCT GTGGGGACCA CGGAAGCCTA CGGGCCAGCC
3121 TCCAAGAATC TGCTCTCCAT CACGGAAGAT CCCGAAGTGG GCACCCCTAC CTATAGCCCG
3181 TCCCTGAAGG AGCTGGACCC CAACCAGCCC CTGGAAAGCA AAGAGCGGAG AGCCAGCGAC
3241 GGGAGCCGGA GCCCCACGCC CAGCCAGAAG CTGGGCAGCG CCTACCTGCC CTCCTATCAC
3301 CACTCCCCAT ACAAGCACGC GCACATCCCG GCGCACGCCC AGCACTACCA GAGCTACATG
3361 CAGCTGATCC AGCAGAAGTC GGCCGTGGAG TACGCGCAAA GCCAGATGAG CCTGGTGAGC
3421 ATGTGCAAGG ACCTGAGCTC TCCCACCCCG TCGGAGCCGC GCATGGAGTG GAAGGTGAAG
3481 ATCCGCAGCG ACGGGACGCG CTACATCACC AAGAGGCCCG TGCGGGACCG CCTGCTGCGG
3541 GAGCGCGCCC TGAAGATCCG GGAAGAGCGC AGCGGCATGA CCACCGACGA CGACGCGGTG
3601 AGCGAGATGA AGATGGGGCG CTACTGGAGC AAGGAGGAGA GGAAGCAGCA CCTGGTGAAG
3661 GCCAAGGAGC AGCGGCGGCG GCGCGAGTTC ATGATGCAGA GCAGGTTGGA TTGTCTCAAG
3721 GAGCAGCAAG CAGCCGATGA CAGGAAGGAG ATGAACATTC TCGAACTGAG CCACAAAAAG
3781 ATGATGAAGA AGAGGAATAA GAAAATCTTC GATAACTGGA TGACGATCCA AGAACTCTTA
3841 ACCCACGGCA CAAAATCCCC GGACGGCACT AGAGTATACA ATTCCTTCCT ATCGGTGACT
3901 ACTGTATAAT TTTCACTTCT GCATTATGTA CATAAAGGAG ACCACTACCA CTGGGGTAGA
3961 AATTCCTGCC TCGTTCAATG CGGCAAGTTT TTGTATATAA GATAAGTACG GTCTTCATGT
4021 TTATAGTCCA AATTTGCAAA CCCTACAACT CTGGGTGTCA TAGGTCTATT TTAAGGGAAG
4081 AGAGAGAAAA ACACCCTTAC TATCTTGGAA GGCAATATTA ACAAACAGAG CTTTTTTCAA
4141 ATAGCAATTG TACTTTTCTA CCTGTACCCT TTTACATAAA GTGTTTAAAT TTCAGAAAGA

4201 TCTTTTATTA AGCATACTTT CACAGAATAA CTTGTTTAAA CTATATTCAT ATAAAAAAGT
4261 TAAACACGCT TTTTTTCCTG CCTAAAACAC AAATACAACT GCCAGTATGT ATTTTTAATG
4321 GAACCCTATT TTATAATGGT ACGTTACTGA ATGTGTTTCA TATGCGTGAC CGTTAAGATA
4381 TTATCATTTA GGTGAAGGTT TCAACTCAAA ACCACCCAAC CCGGTGGTTA ACGATTTAAT
4441 ACATATAACC AAACCGGCAG CGTTTAGAGT TGGGATATAC ATTTAAACAT TTTCCTGGTT
4501 AAGGTTCCCA AGAGAGTGTA AAGGTTTTAG CAGAAAGCAA AATATCTTGC ATCTTTATGG
4561 AAGTTTAAAG CATGTTTGCA AATATTGCAG CCCATTGAAA GAATTTGCAT GTACAGGAAA
4621 GTTGTGGATG GAGACGGTTT GTGGAATTTT AAGTGCTCAT TGTAGTAAAC TTTTGCTTTG
4681 TAGATTTGAA GGTACAGACT TATACAGGCA AGTTCACAAA ATCATGATTA GTTACAAACA
4741 GTAAAATGAA GTTAAAATAA ATTATTATTT TCT

Figure 4

```
   1 MGFELDRFDG DVDPDLKCAL CHKVLEDPLT TPCGHVFCAG CVLPWVVQEG SCPARCRGRL
  61 SAKELNHVLP LKRLILKLDI KCAYATRGCG RVVKLQQLPE HLERCDFAPA RCRHAGCGQV
 121 LLRRDVEAHM RDACDARPVG RCQEGCGLPL THGEQRAGGH CCARALRAHN GALQARLGAL
 181 HKALKKEALR AGKREKSLVA QLAAAQLELQ MTALRYQKKF TEYSARLDSL SRCVAAPPGG
 241 KGEETKSLTL VLHRDSGSLG FNIIGGRPSV DNHDGSSSEG IFVSKIVDSG PAAKEGGLQI
 301 HDRIIEVNGR DLSRATHDQA VEAFKTAKEP IVVQVLRRTP RTKMFTPPSE SQLVDTGTQT
 361 DITFEHIMAL TKMSSPSPPV LDPYLLPEEH PSAHEYYDPN DYIGDIHQEM DREELELEEV
 421 DLYRMNSQDK LGLTVCYRTD DEDDIGIYIS EIDPNSIAAK DGRIREGDRI IQINGIEVQN
 481 REEAVALLTS EENKNFSLLI ARPELQLDEG WMDDDRNDFL DDLHMDMLEE QHHQAMQFTA
 541 SVLQQKKHDE DGGTTDTATI LSNQHEKDSG VGRTDESTRN DESSEQENNG DDATASSNPL
 601 AGQRKLTCSQ DTLGSGDLPF SNESFISADC TDADYLGIPV DECERFRELL ELKCQVKSAT
 661 PYGLYYPSGP LDAGKSDPES VDKELELLNE ELRSIELECL SIVRAHKMQQ LKEQYRESWM
 721 LHNSGFRNYN TSIDVRRHEL SDITELPEKS DKDSSSAYNT GESCRSTPLT LEISPDNSLR
 781 RAAEGISCPS SEGAVGTTEA YGPASKNLLS ITEDPEVGTP TYSPSLKELD PNQPLESKER
 841 RASDGSRSPT PSQKLGSAYL PSYHHSPYKH AHIPAHAQHY QSYMQLIQQK SAVEYAQSQM
 901 SLVSMCKDLS SPTPSEPRME WKVKIRSDGT RYITKRPVRD RLLRERALKI REERSGMTTD
 961 DDAVSEMKMG RYWSKEERKQ HLVKAKEQRR RREFMMQSRL DCLKEQQAAD DRKEMNILEL
1021 SHKKMMKKRN KKIFDNWMTI QELLTHGTKS PDGTRVYNSF LSVTTV
```

Figure 5

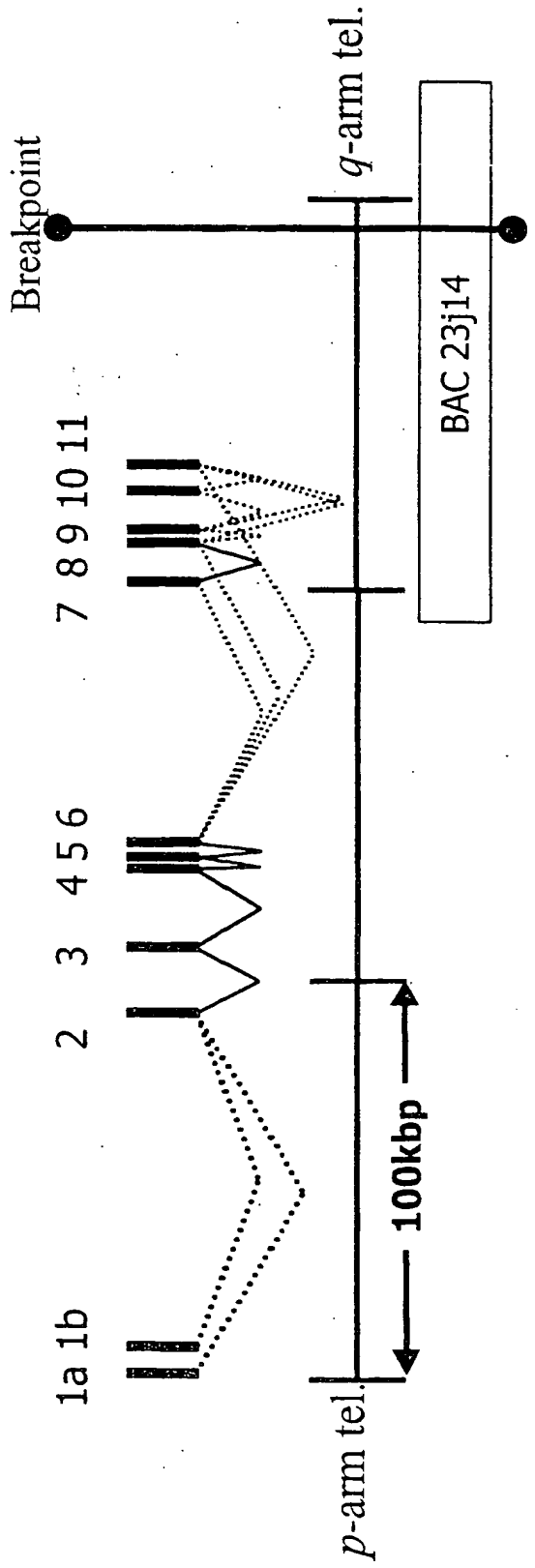

Breakpoint
q-arm tel.
BAC 23j14
7 8 9 10 11
2 3 4 5 6
1a 1b
100kbp
p-arm tel.

Figure 6

1a
atcttcctcctgctctggctgtgtgaagatctgcctccttcctcttcggcttcatgcat
gatcgtaagtttcctgaggcctcctcagccatgcttcctgcatagcctgcagaaat

1b
cccgggtccctcgcaaagccgctgccatcccggagggcccagccagcgggctcccggag
gctggccgggcaggcgtggtgcgcggtaggagctgggcgcgcacggctaccgcgcgtgg
aggagacactgccctgccgcgatgggggcccggggcgctccttcacgccgtaggcaagc
ggggcggcggctgcggtacctgcccaccgggagctttcccttccttctcctgctgctgc
tgctctgcatccagctcgggggaggacagaagaaaaaggag

2-6
These exons have been joined together as they are always spliced in this way.
aatcttttagctgaaaaagtagagcagctgatggaatggagttccagacgctcaatctt
ccgaatgaatggtgataaattccgaaaatttataaaggcaccacctcgaaactattcca
tgattgttatgttcactgctcttcagcctcagcggcagtgttctgtgtgcaggcaagct
aatgaagaatatcaaatactggcgaactcctggcgctattcatctgcttttgtaacaa
gctcttcttcagtatggtggactatgatgaggggacagacgttttcagcagctcaaca
tgaactctgctcctacattcatgcattttcctccaaaaggcagacctaagagagctgat
acttttgacctccaaagaattggatttgcagctgagcaactagcaaagtggattgctga
cagaacggatgttcatattcgggttttcagaccacccaactactctggtaccattgctt
tggccctgttagtgtcgcttgttggaggtttgctttatttgagaaggaacaacttggag
ttcatctataacaagactggttgggccatggtgtctctgtgtatagtctttgctatgac
ttctggccagatgtggaaccatatccgtggacctccatatgctcataagaacccacaca
atggacaagtg

7
agctacattcatgggagcagccaggctcagtttgtggcagaatcacacattattctggt
actga

8
atgccgctatcaccatggggatggttcttctaaatgaagcagcaacttcgaaaggcgat
gttggaaaaagacgga

8+
This is identical to 8 except a cryptic splice acceptor upstream is employed.
tttaaccattctggaacattgtgttcagagccagaaaaattaatagattttattcacat
ctatgtctacggcttccttgacaactactgcagatgccgctatcaccatggggatggtt
cttctaaatgaagcagcaacttcgaaaggcgatgttggaaaaagacgga

9
taatttgcctagtgggattgggcctggtggtcttcttcttcagttttctactttcaata
tttcgttccaagtaccacggctatccttatag

10
tgatctggactttgagtgagaagatgtgatttggaccatggcacttaaaaactctataa
cctcag

11
ctttttaattaaatgaagccaagtgggatttgcataaagtgaatgtttaccatgaagat
aaactgttcctgactttatactattttgaattc

Figure 7

Alternative start exons

1a:

MEWSSRRSIFRMNGDKFRKFIKAPPRNYS..... (encoded by exon 2).

1b:

MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLCIQLGGGQKKKENLLAEKVEQLMEW
SSRRSIFRMNGDKFRKFIKAPPRNYS.....

Transcript options

2-6,7,8,9,10,11
aatcttttagctgaaaaagtagagcagctgatggaatggagttccagacgctcaatctt
ccgaatgaatggtgataaattccgaaaatttataaaggcaccacctcgaaactattcca
tgattgttatgttcactgctcttcagcctcagcggcagtgttctgtgtgcaggcaagct
aatgaagaatatcaaatactggcgaactcctggcgctattcatctgcttttgtaacaa
gctcttcttcagtatggtggactatgatgaggggacagacgttttcagcagctcaaca
tgaactctgctcctacattcatgcattttcctccaaaaggcagacctaagagagctgat
acttttgacctccaaagaattggatttgcagctgagcaactagcaaagtggattgctga
cagaacggatgttcatattcgggttttcagaccacccaactactctggtaccattgctt
tggccctgttagtgtcgcttgttggaggtttgctttatttgagaaggaacaacttggag
ttcatctataacaagactggttgggccatggtgtctctgtgtatagtctttgctatgac
ttctggccagatgtggaaccatatccgtggacctccatatgctcataagaacccacaca
atggacaagtgagctacattcatgggagcagccaggctcagtttgtggcagaatcacac
attattctggtactgaatgccgctatcaccatggggatggttcttctaaatgaagcagc
aacttcgaaaggcgatgttggaaaaagacggataatttgcctagtgggattgggcctgg
tggtcttcttcttcagttttctactttcaatatttcgttccaagtaccacggctatcct
tatagtgatctggactttgagtgagaagatgtgatttggaccatggcacttaaaaactc
tataacctcagcttttaattaaatgaagccaagtgggatttgcataaagtgaatgttt
accatgaagataaactgttcctgactttatactattttgaattc

(MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLCIQLGGGQKKKENLLAEKVEQL)M
EWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTALQPQRQCSVCRQANEEYQILANSW
RYSSAFCNKLFFSMVDYDEGTDVFQQLNMNSAPTFMHFPPKGRPKRADTFDLQRIGFAA
EQLAKWIADRTDVHIRVFRPPNYSGTIALALLVSLVGGLLYLRRNNLEFIYNKTGWAMV
SLCIVFAMTSGQMWNHIRGPPYAHKNPHNGQVSYIHGSSQAQFVAESHIILVLNAAITM
GMVLLNEAATSKGDVGKRRIICLVGLGLVVFFFSFLLSIFRSKYHGYPYSDLDFE

2-6,7,8,9,11

aatcttttagctgaaaaagtagagcagctgatggaatggagttccagacgctcaatctt
ccgaatgaatggtgataaattccgaaaatttataaaggcaccacctcgaaactattcca
tgattgttatgttcactgctcttcagcctcagcggcagtgttctgtgtgcaggcaagct
aatgaagaatatcaaatactggcgaactcctggcgctattcatctgcttttgtaacaa
gctcttcttcagtatggtggactatgatgaggggacagacgtttttcagcagctcaaca
tgaactctgctcctacattcatgcattttcctccaaaaggcagacctaagagagctgat
acttttgacctccaaagaattggatttgcagctgagcaactagcaaagtggattgctga
cagaacggatgttcatattcgggttttcagaccacccaactactctggtaccattgctt
tggccctgttagtgtcgcttgttggaggtttgctttatttgagaaggaacaacttggag
ttcatctataacaagactggttgggccatggtgtctctgtgtatagtctttgctatgac
ttctggccagatgtggaaccatatccgtggacctccatatgctcataagaacccacaca
atggacaagtgagctacattcatgggagcagccaggctcagtttgtggcagaatcacac
attattctggtactgaatgccgctatcaccatggggatggttcttctaaatgaagcagc
aacttcgaaaggcgatgttggaaaaagacggataatttgcctagtgggattgggcctgg
tggtcttcttcttcagttttctactttcaatatttcgttccaagtaccacggctatcct
tatagctttttaattaaatgaagccaagtgggatttgcataaagtgaatgtttaccatg
aagataaactgttcctgactttatactattttgaattc

(MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLCIQLGGGQKKKENLLAEKVEQL)M
EWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTALQPQRQCSVCRQANEEYQILANSW
RYSSAFCNKLFFSMVDYDEGTDVFQQLNMNSAPTFMHFPPKGRPKRADTFDLQRIGFAA
EQLAKWIADRTDVHIRVFRPPNYSGTIALALLVSLVGGLLYLRRNNLEFIYNKTGWAMV
SLCIVFAMTSGQMWNHIRGPPYAHKNPHNGQVSYIHGSSQAQFVAESHIILVLNAAITM
GMVLLNEAATSKGDVGKRRIICLVGLGLVVFFFSFLLSIFRSKYHGYPYSFLIK

2-6,11

aatcttttagctgaaaaagtagagcagctgatggaatggagttccagacgctcaatctt
ccgaatgaatggtgataaattccgaaaatttataaaggcaccacctcgaaactattcca
tgattgttatgttcactgctcttcagcctcagcggcagtgttctgtgtgcaggcaagct
aatgaagaatatcaaatactggcgaactcctggcgctattcatctgcttttgtaacaa
gctcttcttcagtatggtggactatgatgaggggacagacgtttttcagcagctcaaca
tgaactctgctcctacattcatgcattttcctccaaaaggcagacctaagagagctgat
acttttgacctccaaagaattggatttgcagctgagcaactagcaaagtggattgctga
cagaacggatgttcatattcgggttttcagaccacccaactactctggtaccattgctt
tggccctgttagtgtcgcttgttggaggtttgctttatttgagaaggaacaacttggag
ttcatctataacaagactggttgggccatggtgtctctgtgtatagtctttgctatgac
ttctggccagatgtggaaccatatccgtggacctccatatgctcataagaacccacaca
atggacaagtgctttttaattaaatgaagccaagtgggatttgcataaagtgaatgttt
accatgaagataaactgttcctgactttatactattttgaattc

(MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLCIQLGGGQKKKENLLAEKVEQL)M
EWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTALQPQRQCSVCRQANEEYQILANSW
RYSSAFCNKLFFSMVDYDEGTDVFQQLNMNSAPTFMHFPPKGRPKRADTFDLQRIGFAA

EQLAKWIADRTDVHIRVFRPPNYSGTIALALLVSLVGGLLYLRRNNLEFIYNKTGWAMV
SLCIVFAMTSGQMWNHIRGPPYAHKNPHNGQVLFN

2-6,7,8,11
aatcttttagctgaaaaagtagagcagctgatggaatggagttccagacgctcaatctt
ccgaatgaatggtgataaattccgaaaatttataaaggcaccacctcgaaactattcca
tgattgttatgttcactgctcttcagcctcagcggcagtgttctgtgtgcaggcaagct
aatgaagaatatcaaatactggcgaactcctggcgctattcatctgcttttgtaacaa
gctcttcttcagtatggtggactatgatgaggggacagacgttttcagcagctcaaca
tgaactctgctcctacattcatgcattttcctccaaaaggcagacctaagagagctgat
acttttgacctccaaagaattggatttgcagctgagcaactagcaaagtggattgctga
cagaacggatgttcatattcgggttttcagaccacccaactactctggtaccattgctt
tggccctgttagtgtcgcttgttggaggtttgctttatttgagaaggaacaacttggag
ttcatctataacaagactggttgggccatggtgtctctgtgtatagtctttgctatgac
ttctggccagatgtggaaccatatccgtggacctccatatgctcataagaacccacaca
atggacaagtgagctacattcatgggagcagccaggctcagtttgtggcagaatcacac
attattctggtactgaatgccgctatcaccatggggatggttcttctaaatgaagcagc
aacttcgaaaggcgatgttggaaaaagacggacttttttaattaaatgaagccaagtggg
atttgcataaagtgaatgtttaccatgaagataaactgttcctgactttatactatttt
gaattc

(MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLCIQLGGGQKKKENLLAEKVEQL)M
EWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTALQPQRQCSVCRQANEEYQILANSW
RYSSAFCNKLFFSMVDYDEGTDVFQQLNMNSAPTFMHFPPKGRPKRADTFDLQRIGFAA
EQLAKWIADRTDVHIRVFRPPNYSGTIALALLVSLVGGLLYLRRNNLEFIYNKTGWAMV
SLCIVFAMTSGQMWNHIRGPPYAHKNPHNGQVSYIHGSSQAQFVAESHIILVLNAAITM
GMVLLNEAATSKGDVGKRRTF

2-6,8+,9,11
aatcttttagctgaaaaagtagagcagctgatggaatggagttccagacgctcaatctt
ccgaatgaatggtgataaattccgaaaatttataaaggcaccacctcgaaactattcca
tgattgttatgttcactgctcttcagcctcagcggcagtgttctgtgtgcaggcaagct
aatgaagaatatcaaatactggcgaactcctggcgctattcatctgcttttgtaacaa
gctcttcttcagtatggtggactatgatgaggggacagacgttttcagcagctcaaca
tgaactctgctcctacattcatgcattttcctccaaaaggcagacctaagagagctgat
acttttgacctccaaagaattggatttgcagctgagcaactagcaaagtggattgctga
cagaacggatgttcatattcgggttttcagaccacccaactactctggtaccattgctt
tggccctgttagtgtcgcttgttggaggtttgctttatttgagaaggaacaacttggag
ttcatctataacaagactggttgggccatggtgtctctgtgtatagtctttgctatgac
ttctggccagatgtggaaccatatccgtggacctccatatgctcataagaacccacaca
atggacaagtgtttaaccattctggaacattgtgttcagagccagaaaaattaatagat
tttattcacatctatgtctacggcttccttgacaactactgcagatgccgctatcacca
tggggatggttcttctaaatgaagcagcaacttcgaaaggcgatgttggaaaaagacgg
ataatttgcctagtgggattgggcctggtggtcttcttcttcagttttctactttcaat
atttcgttccaagtaccacggctatccttatagcttttttaattaaatgaagccaagtgg

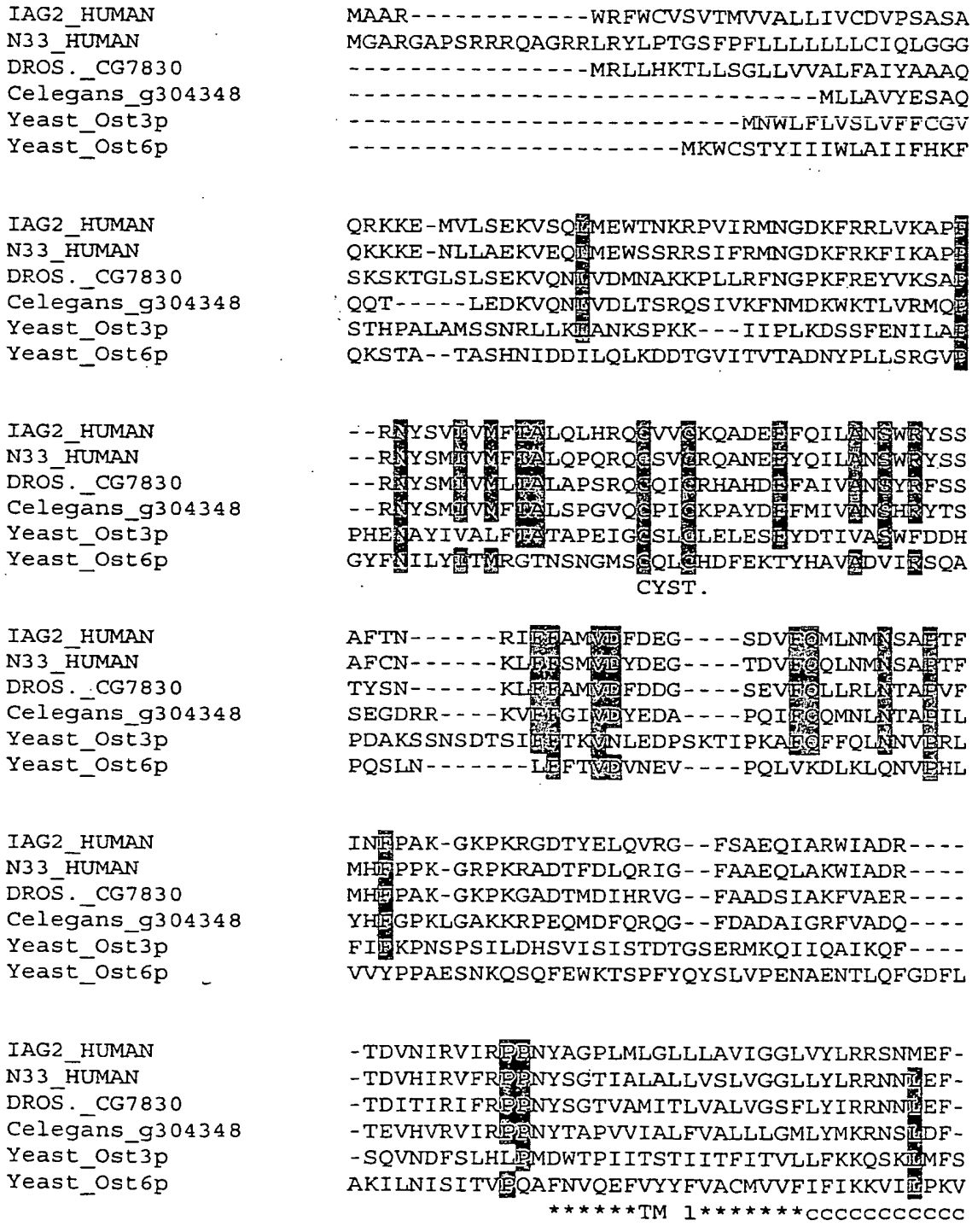
gatttgcataaagtgaatgtttaccatgaagataaactgttcctgactttatactattt
tgaattc

(MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLCIQLGGGQKKKENLLAEKVEQL)M
EWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTALQPQRQCSVCRQANEEYQILANSW
RYSSAFCNKLFFSMVDYDEGTDVFQQLNMNSAPTFMHFPPKGRPKRADTFDLQRIGFAA
EQLAKWIADRTDVHIRVFRPPNYSGTIALALLVSLVGGLLYLRRNNLEFIYNKTGWAMV
SLCIVFAMTSGQMWNHIRGPPYAHKNPHNGQVFNHSGTLCSEPEKLIDFIHIYVYGFLD
NYCRCRYHHGDGSSK

2-6,8+,11
aatcttttagctgaaaaagtagagcagctgatggaatggagttccagacgctcaatctt
ccgaatgaatggtgataaattccgaaaatttataaaggcaccacctcgaaactattcca
tgattgttatgttcactgctcttcagcctcagcggcagtgttctgtgtgcaggcaagct
aatgaagaatatcaaatactggcgaactcctggcgctattcatctgcttttgtaacaa
gctcttcttcagtatggtggactatgatgaggggacagacgtttttcagcagctcaaca
tgaactctgctcctacattcatgcattttcctccaaaaggcagacctaagagagctgat
acttttgacctccaaagaattggatttgcagctgagcaactagcaaagtggattgctga
cagaacggatgttcatattcgggttttcagaccacccaactactctggtaccattgctt
tggccctgttagtgtcgcttgttggaggtttgctttatttgagaaggaacaacttggag
ttcatctataacaagactggttgggccatggtgtctctgtgtatagtctttgctatgac
ttctggccagatgtggaaccatatccgtggacctccatatgctcataagaacccacaca
atggacaagtgtttaaccattctggaacattgtgttcagagccagaaaaattaatagat
tttattcacatctatgtctacggcttccttgacaactactgcagatgccgctatcacca
tggggatggttcttctaaatgaagcagcaacttcgaaaggcgatgttggaaaaagacgg
acttttaattaaatgaagccaagtgggatttgcataaagtgaatgtttaccatgaaga
taaactgttcctgactttatactattttgaattc

(MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLCIQLGGGQKKKENLLAEKVEQL)M
EWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTALQPQRQCSVCRQANEEYQILANSW
RYSSAFCNKLFFSMVDYDEGTDVFQQLNMNSAPTFMHFPPKGRPKRADTFDLQRIGFAA
EQLAKWIADRTDVHIRVFRPPNYSGTIALALLVSLVGGLLYLRRNNLEFIYNKTGWAMV
SLCIVFAMTSGQMWNHIRGPPYAHKNPHNGQVFNHSGTLCSEPEKLIDFIHIYVYGFLD
NYCRCRYHHGDGSSK

Figure 8


IAG2_HUMAN
N33_HUMAN
DROS._CG7830
Celegans_g304348
Yeast_Ost3p
Yeast_Ost6p
CYST.
******TM 1*******ccccccccccc

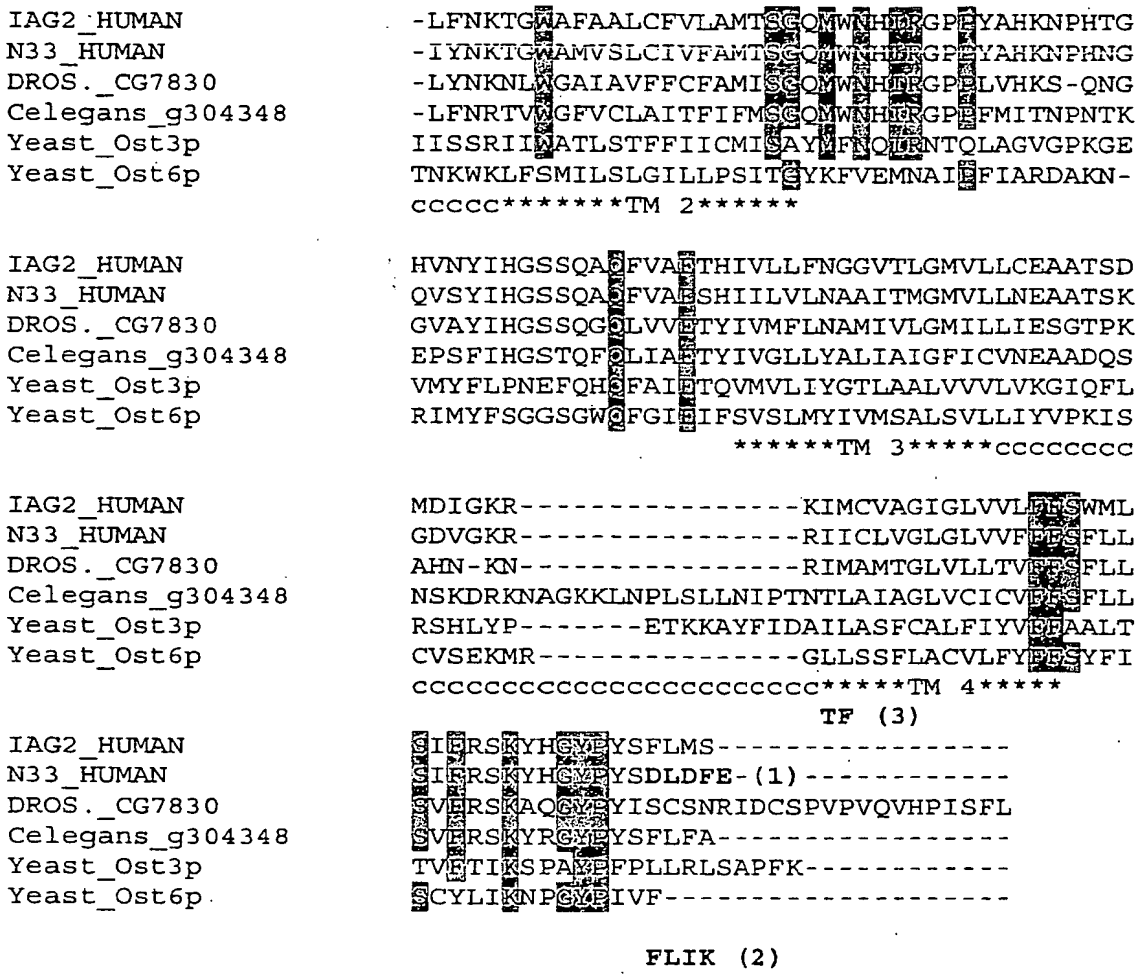

IAG2_HUMAN
N33_HUMAN
DROS._CG7830
Celegans_g304348
Yeast_Ost3p
Yeast_Ost6p
ccccc*******TM 2******
******TM 3*****cccccccc
*****TM 4*****
TF (3)
FLIK (2)

Figure 9

C-termini of N33 splice forms

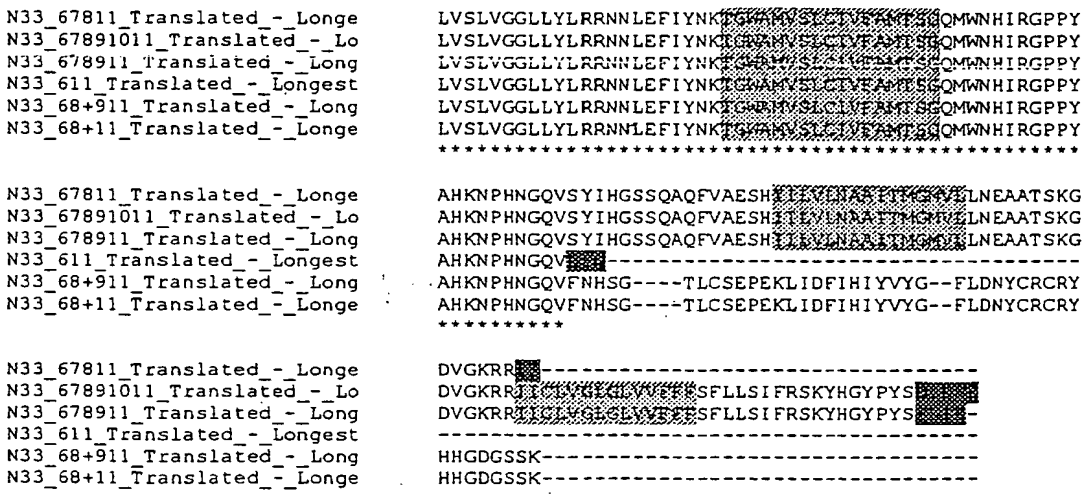


Figure 10

Published GRIK4 nucleic acid sequence (accession NM_014619).

```
   1 atgccccgcg tctcggcgcc tttggtgctg cttcctgcgt ggctcgtgat ggtcgcctgc
  61 agcccgcact ccttgaggat cgctgctatc ttggacgacc ccatggagtg cagcagaggg
 121 gagcggctct ccatcaccct ggccaagaac cgcatcaacc gcgctcctga gaggctgggc
 181 aaggccaagg tcgaagtgga catctttgag cttctcagag acagcgagta cgagactgca
 241 gaaaccatgt gtcagatcct ccccaagggg gtggtcgctg tcctcggacc atcgtccagc
 301 ccagcctcca gctccatcat cagcaacatc tgtggagaga aggaggtccc tcacttcaaa
 361 gtggccccag aggagttcgt caagttccag ttccagagat tcacaaccct gaacctccac
 421 cccagcaaca ctgacatcag cgtggctgta gctgggatcc tgaacttctt caactgcacc
 481 accgcctgcc tcatctgtgc caaagcagaa tgccttttaa acctagagaa gctgctccgg
 541 caattcctta tctccaagga cacgctgtcc gtccgcatgc tggatgacac ccgggacccc
 601 accccgctcc tcaaggagat ccgggacgac aagaccgcca ccatcatcat ccacgccaac
 661 gcctccatgt cccacaccat cctcctgaag gcagccgaac ttgggatggt gtcagcctat
 721 tacacataca tcttcactaa tctggagttc tcactccaga gaacggacag ccttgtggat
 781 gatcgtgtca acatcctggg attttccatt ttcaaccaat cccatgcttt cttccaagag
 841 tttgcccaga gcctcaacca gtcctggcag gagaactgtg accatgtgcc cttcactggg
 901 cctgcgctct cctcggccct gctgtttgat gctgtctatg ctgtggtgac tgcggtgcag
 961 gaactgaacc ggagccaaga gatcggcgtg aagcccttgt cctgcggctc ggcccagatc
1021 tggcagcacg gcaccagcct catgaactac ctgcgcatgg tagaattgga aggtcttacc
1081 ggccacattg aattcaacag caaaggccag aggtccaact acgctttgaa aatcttacag
1141 ttcacaagga atggttttcg gcagatcggc cagtggcacg tggcagaggg cctcagcatg
1201 gacagccacc tctatgcctc caacatctcg gacactctct tcaacaccac cctggtcgtc
1261 accaccatcc tggaaaaccc atatttaatg ctgaagggga accaccagga gatggaaggc
1321 aatgaccgct acgagggctt ctgtgtggac atgctcaagg agctggcaga gatcctccga
1381 ttcaactaca agatccgcct ggttggggat ggcgtgtacg gcgttcccga ggccaacggc
1441 acctggacgg gaatggtcgg ggagctgatc gctaggaaag cagatctggc tgtggcaggc
1501 ctcaccatta cagctgaacg ggagaaggtg attgatttct ctaagccatt catgactctg
1561 ggaattagca ttctttaccg cattcatatg ggacgcaaac ccggctattt ctccttcctg
1621 gacccatttt ctccgggcgt ctggctcttc atgcttctag cctatctggc cgtcagctgt
1681 gtcctcttcc tggtggctcg gttgacgccc tacgagtggt acagcccaca cccatgtgcc
1741 cagggccggt gcaacctcct ggtgaaccag tactccctgg gcaacagcct ctggtttccg
1801 gtcgggggct tcatgcagca gggctccacc atcgcccctc gcgccttatc cacccgctgt
1861 gtcagtggcg tctggtgggc attcacgctg atcatcatct catcctacac ggccaacctg
1921 gcagccttcc tgaccgtgca gcgcatggat gtgcccattg agtcagtgga tgacctggct
1981 gaccagaccg ccattgaata tggcacaatt cacggaggct ccagcatgac cttcttccaa
2041 aattcccgct accagaccta ccaacgcatg tggaattaca tgtattccaa gcagcccagc
2101 gtgttcgtga agagcacaga ggagggaatc gccagggtgt tgaattccaa ctacgccttc
2161 ctcctggaat ccaccatgaa cgagtactat cggcagcgaa actgcaacct cactcagatt
2221 gggggcctgc tggacaccaa gggctatggg attggcatgc cagtcggctc ggttttccgg
2281 gacgagtttg atctggccat tctccagctg caggagaaca accgcctgga gatcctgaag
2341 cgcaaatggt gggaaggagg gaagtgcccc aaggaggaag atcacagagc taaaggcctg
2401 ggaatggaga atattggtgg aatctttgtg gttcttattt gtggcttaat cgtggccatt
2461 tttatggcta tgttggagtt tttatggact ctcagacact cagaagcaac tgaggtgtcc
2521 gtctgccagg agatggtgac cgagctgcgc agcattatcc tgtgtcagga cagtatccac
2581 ccccgccggc ggcgcgccgc agtcccgccg ccccggcccc ccatccccga ggagcgccga
2641 ccgcggggca cggcgacgct cagcaacggg aagctgtgcg gggcagggga gcccgaccag
2701 ctcgcgcaga gactggcgca ggaggccgcc ctggtggccc gcggctgcac gcacatccgc
2761 gtctgccccg agtgccgccg cttccagggc ctgcgggcac ggccgtcgcc cgcccgcagc
2821 gaggagagcc tggagtggga gaaaaccacc aacagcagcg agcccgagta g
```

Figure 11

Published GRIK4 protein sequence (accession NP_055434).

MPRVSAPLVLLPAWLVMVACSPHSLRIAAILDDPMECSRGERLSITLAKNRINRAPERL
GKAKVEVDIFELLRDSEYETAETMCQILPKGVVAVLGPSSSPASSSIISNICGEKEVPH
FKVAPEEFVKFQFQRFTTLNLHPSNTDISVAVAGILNFFNCTTACLICAKAECLLNLEK
LLRQFLISKDTLSVRMLDDTRDPTPLLKEIRDDKTATIIIHANASMSHTILLKAAELGM
VSAYYTYIFTNLEFSLQRTDSLVDDRVNILGFSIFNQSHAFFQEFAQSLNQSWQENCDH
VPFTGPALSSALLFDAVYAVVTAVQELNRSQEIGVKPLSCGSAQIWQHGTSLMNYLRMV
ELEGLTGHIEFNSKGQRSNYALKILQFTRNGFRQIGQWHVAEGLSMDSHLYASNISDTL
FNTTLVVTTILENPYLMLKGNHQEMEGNDRYEGFCVDMLKELAEILRFNYKIRLVGDGV
YGVPEANGTWTGMVGELIARKADLAVAGLTITAEREKVIDFSKPFMTLGISILYRIHMG
RKPGYFSFLDPFSPGVWLFMLLAYLAVSCVLFLVARLTPYEWYSPHPCAQGRCNLLVNQ
YSLGNSLWFPVGGFMQQGSTIAPRALSTRCVSGVWWAFTLIIISSYTANLAAFLTVQRM
DVPIESVDDLADQTAIEYGTIHGGSSMTFFQNSRYQTYQRMWNYMYSKQPSVFVKSTEE
GIARVLNSNYAFLLESTMNEYYRQRNCNLTQIGGLLDTKGYGIGMPVGSVFRDEFDLAI
LQLQENNRLEILKRKWWEGGKCPKEEDHRAKGLGMENIGGIFVVLICGLIVAIFMAMLE
FLWTLRHSEATEVSVCQEMVTELRSIILCQDSIHPRRRRAAVPPPRPPIPEERRPRGTA
TLSNGKLCGAGEPDQLAQRLAQEAALVARGCTHIRVCPECRRFQGLRARPSPARSEESL
EWEKTTNSSEPE

Figure 12

Cytogenetic Position	Description	Breakpoint YAC Clones	Breakpoint BAC Clones (Acc. No.)
2p12	Inversion breakpoint	915_f_7	-
2q32.1	Inversion breakpoint	941_h_12	RP11-358M9 (AC020595)
2q21.3	Translocation breakpoint	766_c_12	RP11-250H22 (AC011996)
11q23.3	Upper insertion breakpoint	936_d_9	RP11-89P5 (AC009641)
11q24.2	Translocation/Insertion breakpoint	749_d_2	RP11-687M24 (AP001007)

Figure 13

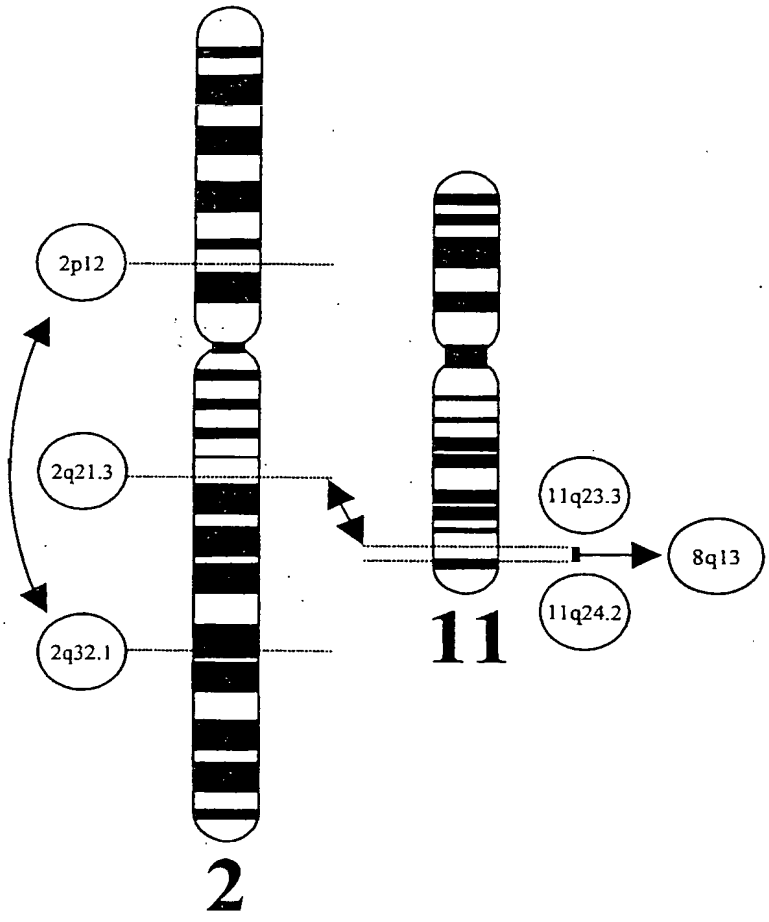


Figure 14

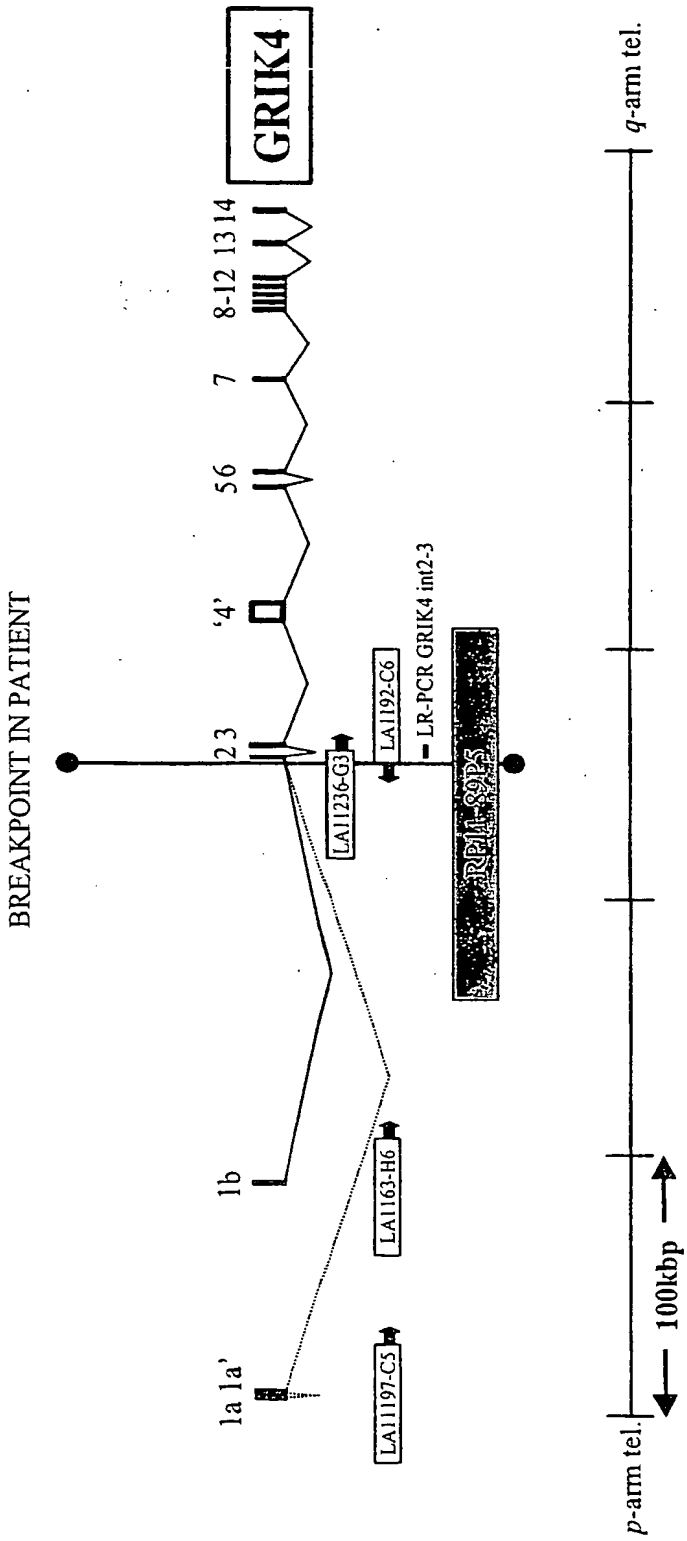

GRIK4
8-12 13 14
7
5 6
'4'
2 3
1b
1a 1a'
BREAKPOINT IN PATIENT
LA11236-G3
LA1192-C6
LR-PCR GRIK4 int2-3
LA1163-H6
LA11197-C5
RP11-89P5
q-arm tel.
p-arm tel.
100kbp

Figure 15

Exon 1a

GCGTGGTAGCATGTGCCTGTAATCCCAGTGCTTTGGGACACCGAGGCAGGAGGATCACT
CGAGCCCAGGAGTGCGAGGCTGCAgtgagttatgatcatac

Exon 1a'

agatttgtcttctctgccagGTGACGCTAGACTTCAGGAAGACCCCCCATTTCTGCTCC
ACTCCTGGGCTTGGAGAAGAGTACAGCTGCTCTTGACTGGTGGGACCTTTTGCTGGCTA
GGGGTGATGGGAGAAGCAAGAGAGGGATCCACACACCTGCGCTTAGCTTTCTATGACCT
GGGCGGATGGAGGCCAAAGgtaaggtgggatgaga
M E A K A

Exon 1b

CCATGAGGATTCATAGAAGATGCCCCGCGTCTCGGCGCCTTTGGTGCTGCTTCCTGCGT
M P R V S A P L V L L P A W
GGCTCGTGATGGTCGCCTGCAGCCCGCACTCCTTGAGGATCGgtaagtgtggcccagct
L V M V A C S P H S L R I A

Exon 2

gaaacccccccccagCTGCTATCTTGGACGACCCCATGGAGTGCAGCAGAGGGGAGCGGC
A I L D D P M E C S R G E R L
TCTCCATCACCCTGGCCAAGAACCGCATCAACCGCGCTCCTGAGAGGCTGGGCAAGGCC
S I T L A K N R I N R A P E R L G K A
AAGGTCGAAGTGGACATCTTTGAGCTTCTCAGAGACAGCGAGTACGAGACTGCAGAAAC
K V E V D I F E L L R D S E Y E T A E T

CAgtacgtagactggg
M

Figure 16
Alternative nucleic acid sequence. Exons 1a-1a'-2-etc.

```
   1 gcgtggtagc atgtgcctgt aatcccagtg ctttgggaca ccgaggcagg aggatcactc
  61 gagcccagga gtgcgaggct gcagtgacgc tagacttcag gaagaccccc catttctgct
 121 ccactcctgg gcttggagaa gagtacagct gctcttgact ggtgggacct tttgctggct
 181 aggggtgatg ggagaagcaa gagagggatc cacacacctg cgcttagctt tctatgacct
 241 gggcggatgg aggccaaagc tgctatcttg gacgacccca tggagtgcag cagaggggag
 301 cggctctcca tcaccctggc caagaaccgc atcaaccgcg ctcctgagag gctgggcaag
 361 gccaaggtcg aagtggacat ctttgagctt ctcagagaca gcgagtacga gactgcagaa
 421 accatgtgtc agatcctccc caaggggtg gtcgctgtcc tcggaccatc gtccagccca
 481 gcctccagct ccatcatcag caacatctgt ggagagaagg aggtccctca cttcaaagtg
 541 gccccagagg agttcgtcaa gttccagttc cagagattca caaccctgaa cctccacccc
 601 agcaacactg acatcagcgt ggctgtagct gggatcctga acttcttcaa ctgcaccacc
 661 gcctgcctca tctgtgccaa agcagaatgc cttttaaacc tagagaagct gctccggcaa
 721 ttccttatct ccaaggacac gctgtccgtc cgcatgctgg atgacacccg ggaccccacc
 781 ccgctcctca aggagatccg ggacgacaag accgccacca tcatcatcca cgccaacgcc
 841 tccatgtccc acaccatcct cctgaaggca gccgaacttg ggatggtgtc agcctattac
 901 acatacatct tcactaatct ggagttctca ctccagagaa cggacagcct tgtggatgat
 961 cgtgtcaaca tcctgggatt ttccattttc aaccaatccc atgctttctt ccaagagttt
1021 gcccagagcc tcaaccagtc ctggcaggag aactgtgacc atgtgccctt cactgggcct
1081 gcgctctcct cggccctgct gtttgatgct gtctatgctg tggtgactgc ggtgcaggaa
1141 ctgaaccgga gccaagagat cggcgtgaag cccttgtcct gcggctcggc ccagatctgg
1201 cagcacggca ccagcctcat gaactacctg cgcatggtag aattggaagg tcttaccggc
1261 cacattgaat tcaacagcaa aggccagagg tccaactacg ctttgaaaat cttacagttc
1321 acaaggaatg gttttcggca gatcggccag tggcacgtgg cagagggcct cagcatggac
1381 agccacctct atgcctccaa catctcggac actctcttca acaccaccct ggtcgtcacc
1441 accatcctgg aaaacccata tttaatgctg aaggggaacc accaggagat ggaaggcaat
1501 gaccgctacg agggcttctg tgtggacatg ctcaaggagc tggcagagat cctccgattc
1561 aactacaaga tccgcctggt tggggatggc gtgtacggcg ttcccgaggc caacggcacc
1621 tggacgggaa tggtcgggga gctgatcgct aggaaagcag atctggctgt ggcaggcctc
1681 accattacag ctgaacggga gaaggtgatt gatttctcta agccattcat gactctggga
1741 attagcattc tttaccgcat tcatatggga cgcaaacccg gctatttctc cttcctggac
1801 ccatttctc cgggcgtctg gctcttcatg cttctagcct atctggccgt cagctgtgtc
1861 ctcttcctgg tggctcggtt gacgccctac gagtggtaca gcccacaccc atgtgcccag
1921 ggccggtgca acctcctggt gaaccagtac tccctgggca acagcctctg gtttccggtc
1981 gggggttca tgcagcaggg ctccaccatc gcccctcgcg ccttatccac ccgctgtgtc
2041 agtggcgtct ggtgggcatt cacgctgatc atcatctcat cctacacggc caacctggca
2101 gccttcctga ccgtgcagcg catggatgtg cccattgagt cagtggatga cctggctgac
2161 cagaccgcca ttgaatatgg cacaattcac ggaggctcca gcatgacctt cttccaaaat
2221 tcccgctacc agacctacca acgcatgtgg aattacatgt attccaagca gcccagcgtg
2281 ttcgtgaaga gcacagagga gggaatcgcc agggtgttga attccaacta cgccttcctc
2341 ctggaatcca ccatgaacga gtactatcgg cagcgaaact gcaacctcac tcagattggg
2401 ggcctgctgg acaccaaggg ctatgggatt ggcatgccag tcggctcggt tttccgggac
2461 gagtttgatc tggccattct ccagctgcag gagaacaacc gcctggagat cctgaagcgc
2521 aaatggtggg aaggagggaa gtgccccaag gaggaagatc acagagctaa aggcctggga
2581 atggagaata ttggtggaat ctttgtggtt cttatttgtg gcttaatcgt ggccattttt
2641 atggctatgt tggagttttt atggactctc agacactcag aagcaactga ggtgtccgtc
2701 tgccaggaga tggtgaccga gctgcgcagc attatcctgt gtcaggacag tatccacccc
2761 cgccggcggc gcgccgcagt cccgccgccc cggccccca tccccgagga gcgccgaccg
2821 cggggcacgg cgacgctcag caacgggaag ctgtgcgggg caggggagcc cgaccagctc
2881 gcgcagagac tggcgcagga ggccgccctg gtggcccgcg gctgcacgca catccgcgtc
2941 tgccccgagt gccgccgctt ccagggcctg cgggcacggc cgtcgcccgc ccgcagcgag
3001 gagagcctgg agtgggagaa aaccaccaac agcagcgagc ccgagtag
```

Figure 17

Complete alternative protein sequence

MEAKAAILDDPMECSRGERLSITLAKNRINRAPERLGKAKVEVDIFELLRDSEYETAET
MCQILPKGVVAVLGPSSSPASSSIISNICGEKEVPHFKVAPEEFVKFQFQRFTTLNLHP
SNTDISVAVAGILNFFNCTTACLICAKAECLLNLEKLLRQFLISKDTLSVRMLDDTRDP
TPLLKEIRDDKTATIIIHANASMSHTILLKAAELGMVSAYYTYIFTNLEFSLQRTDSLV
DDRVNILGFSIFNQSHAFFQEFAQSLNQSWQENCDHVPFTGPALSSALLFDAVYAVVTA
VQELNRSQEIGVKPLSCGSAQIWQHGTSLMNYLRMVELEGLTGHIEFNSKGQRSNYALK
ILQFTRNGFRQIGQWHVAEGLSMDSHLYASNISDTLFNTTLVVTTILENPYLMLKGNHQ
EMEGNDRYEGFCVDMLKELAEILRFNYKIRLVGDGVYGVPEANGTWTGMVGELIARKAD
LAVAGLTITAEREKVIDFSKPFMTLGISILYRIHMGRKPGYFSFLDPFSPGVWLFMLLA
YLAVSCVLFLVARLTPYEWYSPHPCAQGRCNLLVNQYSLGNSLWFPVGGFMQQGSTIAP
RALSTRCVSGVWWAFTLIIISSYTANLAAFLTVQRMDVPIESVDDLADQTAIEYGTIHG
GSSMTFFQNSRYQTYQRMWNYMYSKQPSVFVKSTEEGIARVLNSNYAFLLESTMNEYYR
QRNCNLTQIGGLLDTKGYGIGMPVGSVFRDEFDLAILQLQENNRLEILKRKWWEGGKCP
KEEDHRAKGLGMENIGGIFVVLICGLIVAIFMAMLEFLWTLRHSEATEVSVCQEMVTEL
RSIILCQDSIHPRRRRAAVPPPRPPIPEERRPRGTATLSNGKLCGAGEPDQLAQRLAQE
AALVARGCTHIRVCPECRRFQGLRARPSPARSEESLEWEKTTNSSEPE

Figure 18

NPAS3 (NM_022123) nucleic acid sequence (spliceform 1b-3-4etc)

```
   1 ccacgcgtcc gacgcccccc acccgggagg ggggagagag gcaaaaagta agagaggaaa
  61 aaaaatagca ggaagatggc gcccaccaag cccagctttc agcaggatcc ttccaggcga
 121 gaacgtttac aagcattgag aaaggagaaa tcccgagatg ctgctcgctc ccgccgggga
 181 aaagaaaact ttgagttcta tgaattggcc aagttgttgc ctcttcctgc agccattacc
 241 agccagctcg acaaggcatc catcattcga cttacaatta gctatctgaa aatgagggac
 301 tttgctaacc agggggaccc tccgtggaac ttgcgaatgg aaggccctcc acctaacaca
 361 tcagtaaaag gtgcacagcg aaggagaagc cccagtgcac tagccattga agtatttgaa
 421 gcacatttgg gaagccacat tttgcagtcc ctggatggct ttgtatttgc actaaatcag
 481 gaaggaaaat ttttgtacat ttccgaaaca gtctccatct acctaggcct ctcacaagtg
 541 gagctgacag gcagcagtgt ctttgactat gtccaccccg gagatcacgt ggagatggct
 601 gagcagctgg gcatgaagct ccccccctggg cggggtctcc tgtcacaggg cactgctgag
 661 gacggagcca gctcagcatc ttcctcctct cagtcggaga cccccgagcc agtggagtca
 721 accagcccca gtctgctaac cactgacaac actcttgagc gttccttttt catccgaatg
 781 aaatctactc tgaccaaacg cggtgtgcac atcaaatcat caggatataa ggtgattcac
 841 ataacaggcc ggctacgcct gagagtgtcg ctgtcccacg ggaggaccgt ccccagccaa
 901 atcatgggtc tcgtggttgt tgcgcatgcc ttgcctcccc ctacgatcaa tgaagtcaga
 961 attgactgcc atatgttcgt cactcgagta aatatggacc tcaatatcat ttactgtgaa
1021 aataggatta gtgattatat ggatctgacc cctgtagata tcgtagggaa gagatgctac
1081 cacttcatcc atgctgaaga cgtggagggc atcaggcaca gtcacttgga cttgctgaat
1141 aagggtcagt gtgtgacaaa gtactatcgc tggatgcaga agaacggagg atatatttgg
1201 atacagtcca gtgccaccat agctattaat gccaagaatg caaatgaaaa gaatatcatc
1261 tgggtgaatt accttcttag caatcctgag tacaaggaca cacccatgga catcgcacag
1321 ctcccccatc tgccggagaa aacttccgaa tcctcggaga catccgactc tgagtcagac
1381 tctaaagaca cctcaggtat tacagaggac aacgagaact ccaagtccga cgagaagggg
1441 aaccagtccg agaacagcga agacccggag cccgaccgga agaagtcggg caacgcgtgt
1501 gacaacgaca tgaactgcaa cgacgacggc cacagctcca gtaacccgga cagccgcgac
1561 agcgacgaca gcttcgagca ctcggacttt gagaacccca aggcgggcga ggacggcttc
1621 ggtgctctgg gcgcgatgca gatcaaggtg gagcgctacg tggagagcga gtcggacctg
1681 cggctgcaga actgcgagtc actcacgtcc gacagcgcca aggactcgga cagcgcaggc
1741 gaggcgggcg cgcaggcctc cagcaagcac cagaagcgca agaaaaggcg gaaacggcaa
1801 aagggcggca gcgccagccg ccggcgcctg tccagcgcgt cgagcccagg cggcctggac
1861 gcgggcctgg tggagccccc gcggctgctg tcctcccccа acagtgcctc ggtgctcaag
1921 atcaagacgg agatctcaga acccatcaat ttcgacaatg acagcagcat ctggaactac
1981 ccgcccaacc gggagatctc caggaacgag tccccctaca gcatgaccaa gccccccagc
2041 tctgagcact tcccgtcccc gcagggcggc ggcggtgggg gtggcggtgg cgggggggctg
2101 cacgtggcca ttcccgactc ggtcctcacc ccgcccggcg ccgacggcgc ggccgcccgc
2161 aagactcagt tcggcgcctc ggccaccgcg gccctggccc ccgtcgcctc cgacccgctg
2221 tcacccccgc tctcggcgtc cccgcgggac aagcaccccg ggaacggcgg cgggggcggg
2281 ggcgggggcg gcggcgcggg gggcggcggc cccagcgcgt ccaactcctt gctgtacact
2341 ggggacctgg aggcgctgca gaggttgcag gcgggcaacg tcgtgctccc gctggtgcac
2401 agggtgaccg ggaccctggc cgccaccagc acggccgcgc agagggtcta caccacgggc
2461 accatccgct acgcgcccgc cgaggtgacc ctggccatgc agagcaacct gctgcccaac
2521 gcgcacgctg ttaacttcgt ggacgttaac agccccggct ttggcctcga ccccaagacg
2581 cccatggaga tgctctacca ccacgtgcac cggctcaaca tgtcaggacc gttcggcggc
2641 gcagtgagcg cagctagcct gacgcagatg cccgccggca acgtgttcac cacggccgag
2701 ggactcttct ccacgctgcc cttccccgtc tacagcaacg gcatccacgc ggcacagact
2761 ctggagcgca aggaggactg aggcgccgcc cgtcctgggc ccggccaggc cccgcttgga
2821 ggaggcatcg tcggcatttt cgtttagacc tttaattcta gcactttgaa ttcgagcagg
```

```
2881 tcagcgtctt ctctcgccac gacggtcccc attccacccc ctctttcttt cacctgactt
2941 attctttcgt gtaaagatat gtttattttt tgccttcaga gggtcagacg accagttgcc
3001 tgccgttttg tcttcttcta aggtgtgtgt tgggttgttt tgctttcctt tgcatcttta
3061 ttaagatgtc tttcatgtgt atatgcctct gccatagaat actcagtctt gtggtcaaga
3121 gagttctcaa gtgacaacca ttggggtttc ttcataaaga tcttgatatg atcaagatgg
3181 aaagagacaa gcataaacaa tgtgccctgt ttgactaagt caaatgaaat agggtggttt
3241 ttgtttctgt tcctaattcc tttaaaaaat agggggaata gtattttaga attttatgca
3301 gaatttaatt ctctttttac ggttaagatt ttaagatttt cttacttgca cataaaaata
3361 atttgggttc ttaaacttaa tttctggcct gtgactagaa tgtttaaaaa aaaaaaaaac
3421 cctcgtgc
```

Figure 19

NPAS3 protein sequence (spliceform 1b-3-4etc.)

MAPTKPSFQQDPSRRERLQALRKEKSRDAARSRRGKENFEFYELAKLLPLPAAITSQLD
KASIIRLTISYLKMRDFANQGDPPWNLRMEGPPPNTSVKGAQRRRSPSALAIEVFEAHL
GSHILQSLDGFVFALNQEGKFLYISETVSIYLGLSQVELTGSSVFDYVHPGDHVEMAEQ
LGMKLPPGRGLLSQGTAEDGASSASSSSQSETPEPVESTSPSLLTTDNTLERSFFIRMK
STLTKRGVHIKSSGYKVIHITGRLRLRVSLSHGRTVPSQIMGLVVVAHALPPPTINEVR
IDCHMFVTRVNMDLNIIYCENRISDYMDLTPVDIVGKRCYHFIHAEDVEGIRHSHLDLL
NKGQCVTKYYRWMQKNGGYIWIQSSATIAINAKNANEKNIIWVNYLLSNPEYKDTPMDI
AQLPHLPEKTSESSETSDSESDSKDTSGITEDNENSKSDEKGNQSENSEDPEPDRKKSG
NACDNDMNCNDDGHSSSNPDSRDSDDSFEHSDFENPKAGEDGFGALGAMQIKVERYVES
ESDLRLQNCESLTSDSAKDSDSAGEAGAQASSKHQKRKKPAKRQKGGSASRRRLSSASS
PGGLDAGLVEPPRLLSSPNSASVLKIKTEISEPINFDNDSSIWNYPPNREISRNESPYS
MTKPPSSEHFPSPQGGGGGGGGGGGLHVAIPDSVLTPPGADGAAARKTQFGASATAALA
PVASDPLSPPLSASPRDKHPGNGGGGGGGGGAGGGGPSASNSLLYTGDLEALQRLQAG
NVVLPLVHRVTGTLAATSTAAQRVYTTGTIRYAPAEVTLAMQSNLLPNAHAVNFVDVNS
PGFGLDPKTPMEMLYHHVHRLNMSGPFGGAVSAASLTQMPAGNVFTTAEGLFSTLPFPV
YSNGIHAAQTLERKED
```
```

Figure 20

NPAS3 nucleic acid sequence (spliceform incorporating exons 1a-2-3-4etc) similar to mouse cDNA with accession number NM_013780)

```
   1 ATGGGGAGGG CCGGCGCCGC GGCCAACGGC ACCCCGCAGA ACGTCCAGGG CATCACCTCC
  61 TACCAGCAGC GAATAACTGC CCAGCATCCT CTGCCCAACC AATCAGAATG TAGGAAAATC
 121 TACAGATATG ACGGAATCTA CTGTGAATCT ACCTACCAGA ATTTACAAGC ATTGAGAAAG
 181 GAGAAATCCC GAGATGCTGC TCGCTCCCGC CGGGGAAAAG AAAACTTTGA GTTCTATGAA
 241 TTGGCCAAGT TGTTGCCTCT TCCTGCAGCC ATTACCAGCC AGCTCGACAA GGCATCCATC
 301 ATTCGACTTA CAATTAGCTA TCTGAAAATG AGGGACTTTG CTAACCAGGG GGACCCTCCG
 361 TGGAACTTGC GAATGGAAGG CCCTCCACCT AACACATCAG TAAAAGGTGC ACAGCGAAGG
 421 AGAAGCCCCA GTGCACTAGC CATTGAAGTA TTTGAAGCAC ATTTGGGAAG CCACATTTTG
 481 CAGTCCCTGG ATGGCTTTGT ATTTGCACTA AATCAGGAAG GAAAATTTTT GTACATTTCC
 541 GAAACAGTCT CCATCTACCT AGGCCTCTCA CAAGTGGAGC TGACAGGCAG CAGTGTCTTT
 601 GACTATGTCC ACCCCGGAGA TCACGTGGAG ATGGCTGAGC AGCTGGGCAT GAAGCTCCCC
 661 CCTGGGCGGG GTCTCCTGTC ACAGGGCACT GCTGAGGACG GAGCCAGCTC AGCATCTTCC
 721 TCCTCTCAGT CGGAGACCCC CGAGCCAGTG GAGTCAACCA GCCCCAGTCT GCTAACCACT
 781 GACAACACTC TTGAGCGTTC CTTTTTCATC CGAATGAAAT CTACTCTGAC CAAACGCGGT
 841 GTGCACATCA AATCATCAGG ATATAAGGTG ATTCACATAA CAGGCCGGCT ACGCCTGAGA
 901 GTGTCGCTGT CCCACGGGAG GACCGTCCCC AGCCAAATCA TGGGTCTCGT GGTTGTTGCG
 961 CATGCCTTGC CTCCCCCTAC GATCAATGAA GTCAGAATTG ACTGCCATAT GTTCGTCACT
1021 CGAGTAAATA TGGACCTCAA TATCATTTAC TGTGAAAATA GGATTAGTGA TTATATGGAT
1081 CTGACCCCTG TAGATATCGT AGGGAAGAGA TGCTACCACT TCATCCATGC TGAAGACGTG
1141 GAGGGCATCA GGCACAGTCA CTTGGACTTG CTGAATAAGG GTCAGTGTGT GACAAAGTAC
1201 TATCGCTGGA TGCAGAAGAA CGGAGGATAT ATTTGGATAC AGTCCAGTGC CACCATAGCT
1261 ATTAATGCCA AGAATGCAAA TGAAAAGAAT ATCATCTGGG TGAATTACCT TCTTAGCAAT
1321 CCTGAGTACA AGGACACACC CATGGACATC GCACAGCTCC CCCATCTGCC GGAGAAAACT
1381 TCCGAATCCT CGGAGACATC CGACTCTGAG TCAGACTCTA AAGACACCTC AGGTATTACA
1441 GAGGACAACG AGAACTCCAA GTCCGACGAG AAGGGGAACC AGTCCGAGAA CAGCGAAGAC
1501 CCGGAGCCCG ACCGGAAGAA GTCGGGCAAC GCGTGTGACA ACGACATGAA CTGCAACGAC
1561 GACGGCCACA GCTCCAGTAA CCCGGACAGC CGCGACAGCG ACGACAGCTT CGAGCACTCG
1621 GACTTTGAGA ACCCCAAGGC GGGCGAGGAC GGCTTCGGTG CTCTGGGCGC GATGCAGATC
1681 AAGGTGGAGC GCTACGTGGA GAGCGAGTCG GACCTGCGGC TGCAGAACTG CGAGTCACTC
1741 ACGTCCGACA GCGCCAAGGA CTCGGACAGC GCAGGCGAGG CGGGCGCGCA GGCCTCCAGC
1801 AAGCACCAGA AGCGCAAGAA AAGGCGGAAA CGGCAAAAGG GCGGCAGCGC CAGCCGCCGG
1861 CGCCTGTCCA GCGCGTCGAG CCCAGGCGGC CTGGACGCGG GCCTGGTGGA GCCCCCGCGG
1921 CTGCTGTCCT CCCCCAACAG TGCCTCGGTG CTCAAGATCA AGACGGAGAT CTCAGAACCC
1981 ATCAATTTCG ACAATGACAG CAGCATCTGG AACTACCCGC CCAACCGGGA GATCTCCAGG
2041 AACGAGTCCC CCTACAGCAT GACCAAGCCC CCCAGCTCTG AGCACTTCCC GTCCCCGCAG
2101 GGCGGCGGCG GTGGGGGTGG CGGTGGCGGG GGGCTGCACG TGGCCATTCC CGACTCGGTC
2161 CTCACCCCGC CCGGCGCCGA CGGCGCGGCC GCCCGCAAGA CTCAGTTCGG CGCCTCGGCC
2221 ACCGCGGCCC TGGCCCCCGT CGCCTCCGAC CCGCTGTCAC CCCCGCTCTC GGCGTCCCCG
2281 CGGGACAAGC ACCCCGGGAA CGGCGGCGGG GGCGGGGGCG GGGGCGGCGG CGCGGGGGGC
2341 GGCGGCCCCA GCGCGTCCAA CTCCTTGCTG TACACTGGGG ACCTGGAGGC GCTGCAGAGG
2401 TTGCAGGCGG GCAACGTCGT GCTCCCGCTG GTGCACAGGG TGACCGGGAC CCTGGCCGCC
2461 ACCAGCACGG CCGCGCAGAG GGTCTACACC ACGGGCACCA TCCGCTACGC GCCCGCCGAG
2521 GTGACCCTGG CCATGCAGAG CAACCTGCTG CCCAACGCGC ACGCTGTTAA CTTCGTGGAC
2581 GTTAACAGCC CCGGCTTTGG CCTCGACCCC AAGACGCCCA TGGAGATGCT CTACCACCAC
2641 GTGCACCGGC TCAACATGTC AGGACCGTTC GGCGGCGCAG TGAGCGCAGC TAGCCTGACG
2701 CAGATGCCCG CCGGCAACGT GTTCACCACG GCCGAGGGAC TCTTCTCCAC GCTGCCCTTC
2761 CCCGTCTACA GCAACGGCAT CCACGCGGCA CAGACTCTGG AGCGCAAGGA GGACTGAGGC
2821 GCCGCCCGTC CTGGGCCCGG CCAGGCCCCG CTTGGAGGAG GCATCGTCGG CATTTTCGTT
```

2881 TAGACCTTTA ATTCTAGCAC TTTGAATTCG AGCAGGTCAG CGTCTTCTCT CGCCACGACG
2941 GTCCCCATTC CACCCCCTCT T

Figure 21

NPAS3 protein sequence of spliceform incorporating exons 1a-2-3-4etc.

MGRAGAAANGTPQNVQGTTSVQQRLTAQHPLPNQSTGRKIYRYDGIYGESTMQNLQALR
KEKSRDAARSRRGKENFEFYELAKLLPLPAAITSQLDKASIIRLTISYLKMRDFANQGD
PPWNLRMEGPPPNTSVKGAQRRRSPSALAIEVFEAHLGSHILQSLDGFVFALNQEGKFL
YISETVSIYLGLSQVELTGSSVFDYVHPGDHVEMAEQLGMKLPPGRGLLSQGTAEDGAS
SASSSSQSETPEPVESTSPSLLTTDNTLERSFFIRMKSTLTKRGVHIKSSGYKVIHITG
RLRLRVSLSHGRTVPSQIMGLVVVAHALPPPTINEVRIDCHMFVTRVNMDLNIIYCENR
ISDYMDLTPVDIVGKRCYHFIHAEDVEGIRHSHLDLLNKGQCVTKYYRWMQKNGGYIWI
QSSATIAINAKNANEKNIIWVNYLLSNPEYKDTPMDIAQLPHLPEKTSESSETSDSESD
SKDTSGITEDNENSKSDEKGNQSENSEDPEPDRKKSGNACDNDMNCNDDGHSSSNPDSR
DSDDSFEHSDFENPKAGEDGFGALGAMQIKVERYVESESDLRLQNCESLTSDSAKDSDS
AGEAGAQASSKHQKRKKRRKRQKGGSASRRRLSSASSPGGLDAGLVEPPRLLSSPNSAS
VLKIKTEISEPINFDNDSSIWNYPPNREISRNESPYSMTKPPSSEHFPSPQGGGGGGGG
GGGLHVAIPDSVLTPPGADGAAARKTQFGASATAALAPVASDPLSPPLSASPRDKHPGN
GGGGGGGGGGAGGGGPSASNSLLYTGDLEALQRLQAGNVVLPLVHRVTGTLAATSTAAQ
RVYTTGTIRYAPAEVTLAMQSNLLPNAHAVNFVDVNSPGFGLDPKTPMEMLYHHVHRLN
MSGPFGGAVSAASLTQMPAGNVFTTAEGLFSTLPFPVYSNGIHAAQTLERKED

Figure 22

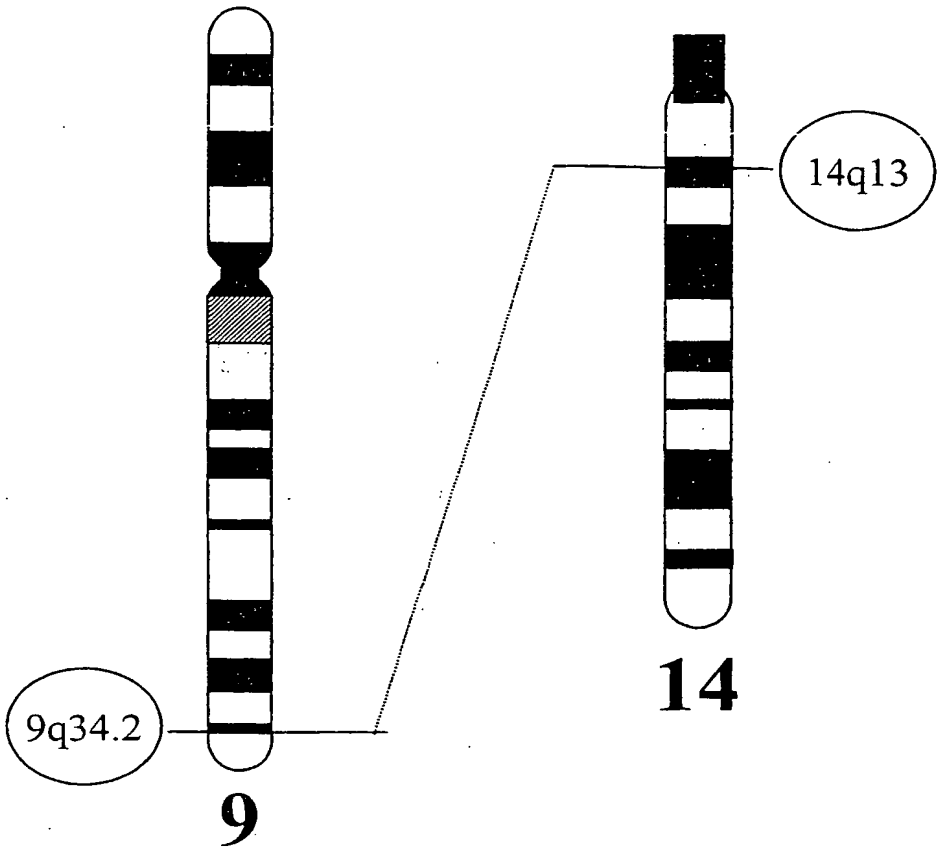


Figure 23

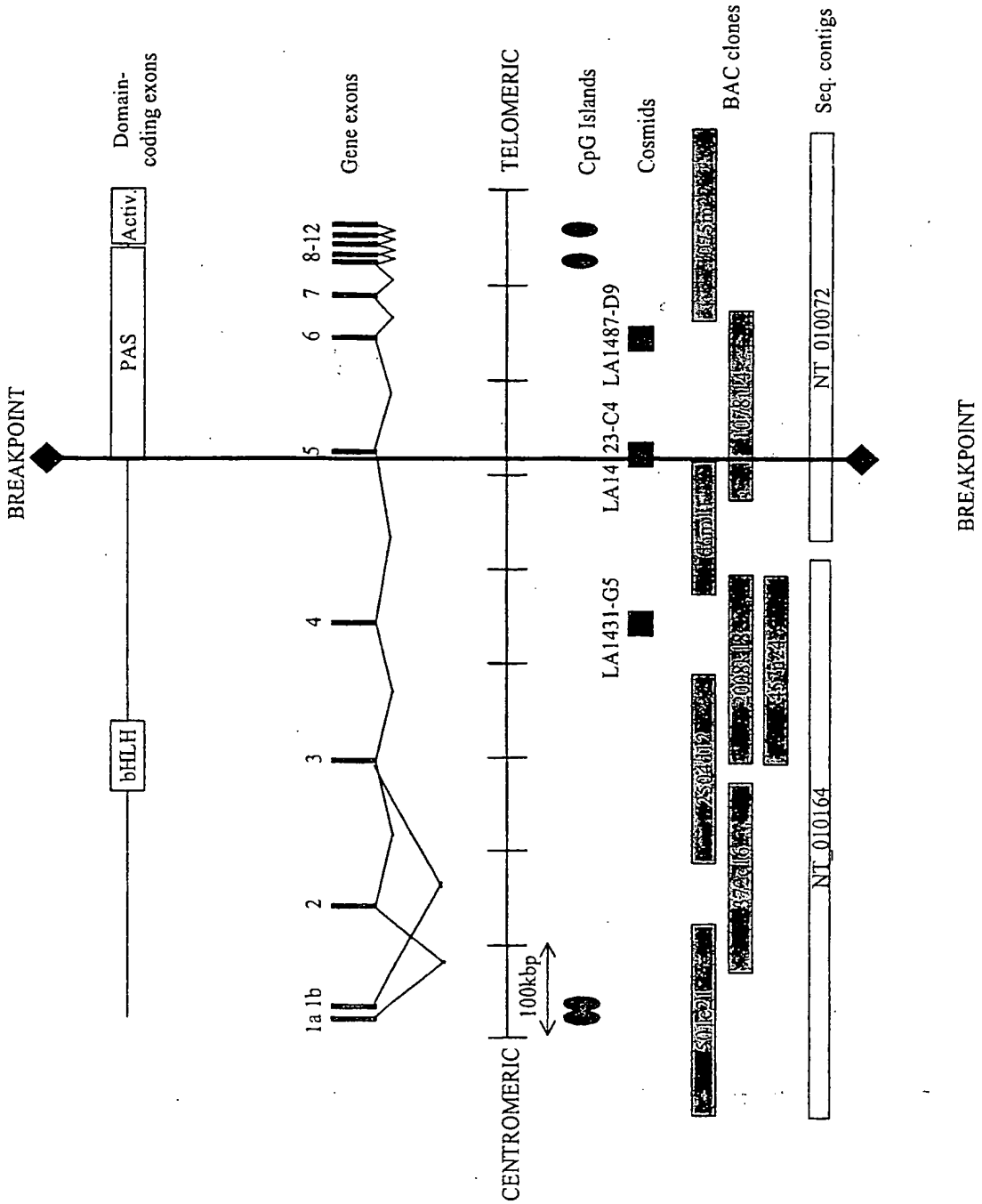

BREAKPOINT
PAS
Activ.
bHLH
Domain-coding exons
1a 1b
2
3
4
5
6
7
8-12
Gene exons
CENTROMERIC
100kbp
TELOMERIC
LA1431-G5
LA14
23-C4
LA1487-D9
CpG Islands
Cosmids
BAC clones
NT_010164
NT_010072
Seq. contigs
BREAKPOINT

Figure 24

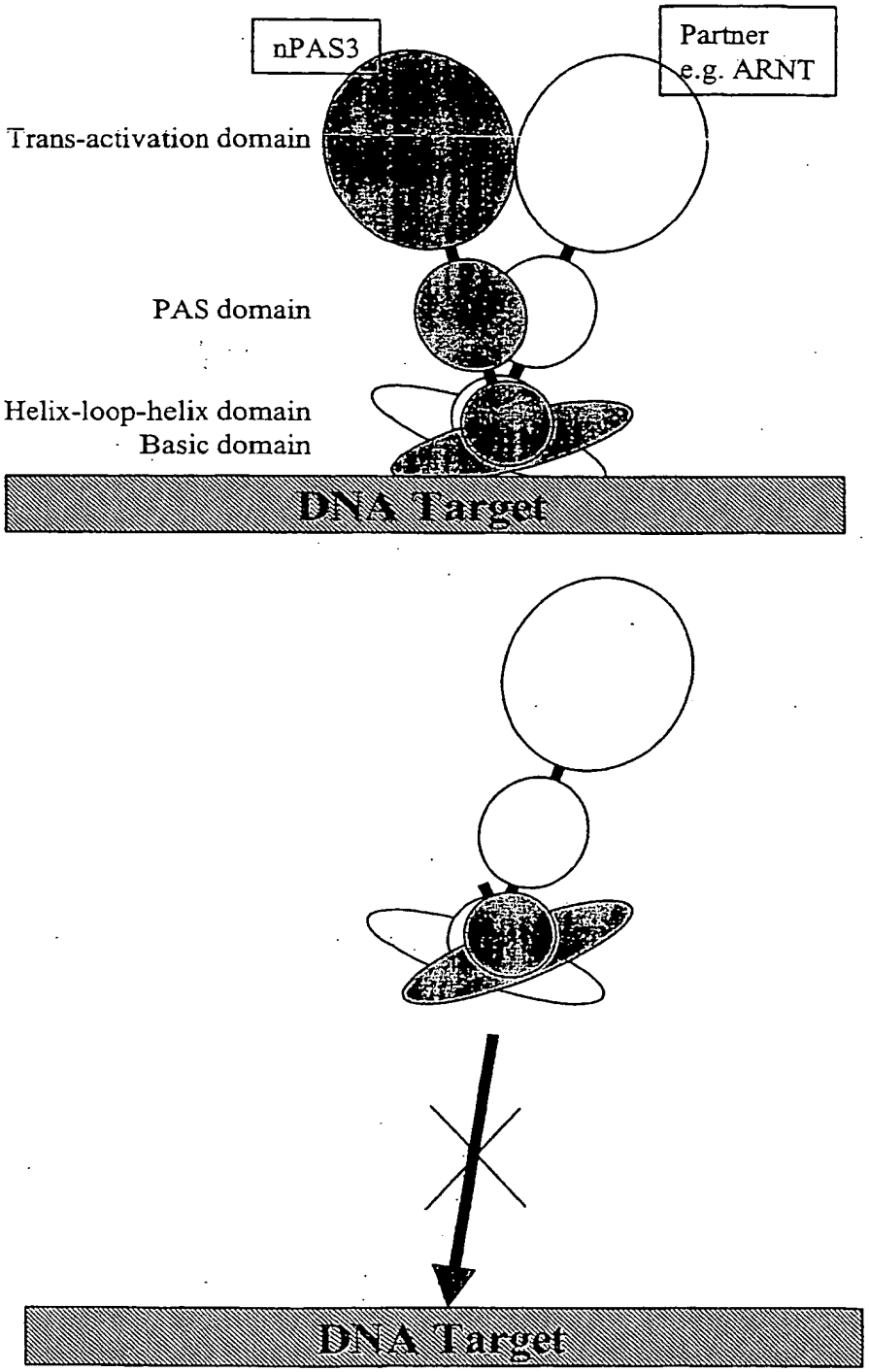

nPAS3
Partner
e.g. ARNT
Trans-activation domain
PAS domain
Helix-loop-helix domain
Basic domain
DNA Target
DNA Target

Figure 25

PDE4B1 (acc. L20966) Nucleic acid sequence

```
   1 gcggccgcgg cggtgcagca gaggcgcctc gggcaggagg agggcggctt ctgcgagggc
  61 agcctgaggt attaaaaagt gtcagcaaac tgcattgaat aacagacatc ctaagagggg
 121 atattttcca cctctataat gaagaaaagc aggagtgtga tgacggtgat ggctgatgat
 181 aatgttaaag attattttga atgtagcttg agtaaatcct acagttcttc cagtaacaca
 241 cttgggatcg acctctggag agggagaagg tgttgctcag gaaacttaca gttaccacca
 301 ctgtctcaaa gacagagtga aagggcaagg actcctgagg gagatggtat ttccaggccg
 361 accacactgc ctttgacaac gcttccaagc attgctatta caactgtaag ccaggagtgc
 421 tttgatgtgg aaaatggccc ttccccaggt cggagtccac tggatcccca ggccagctct
 481 tccgctgggc tggtacttca cgccaccttt cctgggcaca gccagcgcag agagtcattt
 541 ctctacagat cagacagcga ctatgacttg tcaccaaagg cgatgtcgag aaactcttct
 601 cttccaagcg agcaacacgg cgatgacttg attgtaactc cttttgccca ggtccttgcc
 661 agcttgcgaa gtgtgagaaa caacttcact atactgacaa accttcatgg tacatctaac
 721 aagaggtccc cagctgctag tcagcctcct gtctccagag tcaacccaca agaagaatct
 781 tatcaaaaat tagcaatgga aacgctggag gaattagact ggtgtttaga ccagctagag
 841 accatacaga cctaccggtc tgtcagtgag atggcttcta acaagttcaa aagaatgctg
 901 aaccgggagc tgacacacct ctcagagatg agccgatcag ggaaccaggt gtctgaatac
 961 atttcaaata ctttcttaga caagcagaat gatgtggaga tcccatctcc tacccagaaa
1021 gacagggaga aaaagaaaaa gcagcagctc atgacccaga taagtggagt gaagaaatta
1081 atgcatagtt caagcctaaa caatacaagc atctcacgct ttggagtcaa cactgaaaat
1141 gaagatcacc tggccaagga gctggaagac ctgaacaaat ggggtcttaa catctttaat
1201 gtggctggat attctcacaa tagaccccta acatgcatca tgtatgctat attccaggaa
1261 agagacctcc taaagacatt cagaatctca tctgacacat ttataaccta catgatgact
1321 ttagaagacc attaccattc tgacgtggca tatcacaaca gcctgcacgc tgctgatgta
1381 gcccagtcga cccatgttct cctttctaca ccagcattag acgctgtctt cacagatttg
1441 gagatcctgg ctgccatttt tgcagctgcc atccatgacg ttgatcatcc tggagtctcc
1501 aatcagtttc tcatcaacac aaattcagaa cttgctttga tgtataatga tgaatctgtg
1561 ttggaaaatc atcaccttgc tgtgggtttc aaactgctca aagaagaaca ctgtgacatc
1621 ttcatgaatc tcaccaagaa gcagcgtcag acactcagga agatggttat tgacatggtg
1681 ttagcaactg atatgtctaa acatatgagc ctgctggcag acctgaagac aatggtagaa
1741 acgaagaaag ttacaagttc aggcgttctt ctcctagaca actataccga tcgcattcag
1801 gtccttcgca acatggtaca ctgtgcagac ctgagcaacc ccaccaagtc cttggaattg
1861 tatcggcaat ggacagaccg catcatggag gaatttttcc agcagggaga caaagagcgg
1921 gagaggggaa tggaaattag cccaatgtgt gataaacaca cagcttctgt ggaaaaatcc
1981 caggttggtt tcatcgacta cattgtccat ccattgtggg agacatgggc agatttggta
2041 cagcctgatg ctcaggacat tctcgatacc ttagaagata acaggaactg gtatcagagc
2101 atgatacctc aaagtccctc accaccactg gacgagcaga acagggactg ccagggtctg
2161 atggagaagt ttcagtttga actgactctc gatgaggaag attctgaagg acctgagaag
2221 gagggagagg gacacagcta tttcagcagc acaaagacgc tttgtgtgat tgatccagaa
2281 aacagagatt ccctgggaga gactgacata gacattgcaa cagaagacaa gtcccccgtg
2341 gatacataat ccccctctcc ctgtggagat gaacattcta tccttgatga gcatgccagc
2401 tatgtggtag ggccagccca ccatgggggc caagacctgc acaggacaag ggccacctgg
2461 cctttcagtt acttgagttt ggagtcagaa agcaagacca ggaagcaaat agcagctcag
2521 gaaatcccac ggttgacttg ccttgatggc aagcttggtg gagagggctg aagctgttgc
2581 tgggggccga ttctgatcaa gacacatggc ttgaaaatgg aagacacaaa actgagagat
2641 cattctgcac taagtttcgg gaacttatcc ccgacagtga ctgaactcac tgactaataa
2701 cttcatttat gaatcttctc acttgtccct ttgtctgcca acctgtgtgc ctttttgta
2761 aaacattttc atgtctttaa aatgcctgtt gaatacctgg agtttagtat caacttctac
2821 acagataagc tttcaaagtt gacaaacttt tttgactctt tctggaaaag ggaaagaaaa
2881 tagtcttcct tcttttcttgg gcaatatcct tcactttact acagttactt ttgcaaacag
2941 acagaaagga tacacttcta accacatttt acttccttcc cctgttgtcc agtccaactc
3001 cacagtcact cttaaaactt ctctctgttt gcctgcctcc aacagtactt ttaacttttt
```

3061 gctgtaaaca gaataaaatt gaacaaatta gggggtagaa aggagcagtg gtgtcgttca
3121 ccgtgagagt ctgcatagaa ctcagcagtg tgccctgctg tgtcttggac cctgcaatgc
3181 ggccgc

Figure 26

PDE4B1 Protein sequence

MKKSRSVMTVMADDNVKDYFECSLSKSYSSSSNTLGIDFWRGRRCGSCNLQLPPSSQRG
SFRARTPEGDGISRPTTLPLTTLPSIATTIVSQECFDVTNGPSPGRSFLDPQASSSAGT
VLTATFPGHSQRRESTLYRSDSDYDLSPKAMSRNSSLPSEQHGDDLMVTPFAQVLASLR
SVRNNFTILTNLHGISNKRSPAASQPPVSRVNPQEESYQKLAMETLEELDWCLDQLETI
QTYRSVSEMASNKFKRMLNRELTHLSEMSRSGNQVSEYISNTFLDKQNDVEIPSPTQKD
REKKKKQQLMTQISGVKKLMHSSSLNNTSISRFGVNTENEDHLAKELEDLNKWGLNIFN
VAGYSHNRPLTCIMYAIFQERDLLKTFRISSDTFITYMMTLEDHYHSDVAYHNSLHAAD
VAQSTHVLLSTPALDAVFTDLEILAAIFAAAIHDVDHPGVSNQFLINTNSELALMYNDE
SVLENHHLAVGFKLLQEEHCDIFMNLTKKQRQTLRKMVIDMVLATDMSKHMSLLADLKT
MVETKKVTSSGVLLLDNYTDRIQVLRNMVHCADLSNPTKSLELYRQWTDRIMEEFFQQG
DKERERGMEISPMCDKHTASVEKSQVGFIDYIVHPLWETWADLVQPDAQDILDTLEDNR
NWYQSMIPQSPSPPLDEQNRDCQGLMEKFQFELTLDEEDSEGPEKEGEGHSYFSSTKTL
CVIDPENRDSLGETDIDIATEDKSPVDT

Figure 27

PDE4B3 (acc. U85048) Nucleic acid sequence

```
   1 atgacagcaa aagattcttc aaaggaactt actgcttctg aacctgaggt ttgcataaag
  61 actttcaagg agcaaatgca tttagaactt gagcttccga gattaccagg aaacagacct
 121 acatctccta aaatttctcc acgcagttca ccaaggaact caccatgctt tttcagaaag
 181 ttactggtga ataaaagcat tcggcagcgt cgtcgcttca ctgtggctca tacatgcttt
 241 gatgtggaaa atggcccttc cccaggtcgg agtccactgg atccccaggc cagctcttcc
 301 gctgggctgg tacttcacgc cacctttcct gggcacagcc agcgcagaga gtcatttctc
 361 tacagatcag acagcgacta tgacttgtca ccaaaggcga tgtcgagaaa ctcttctctt
 421 ccaagcgagc aacacggcga tgacttgatt gtaactcctt ttgcccaggt ccttgccagc
 481 ttgcgaagtg tgagaaacaa cttcactata ctgacaaacc ttcatggtac atctaacaag
 541 aggtccccag ctgctagtca gcctcctgtc tccagagtca acccacaaga agaatcttat
 601 caaaaattag caatggaaac gctggaggaa ttagactggt gtttagacca gctagagacc
 661 atacagacct accggtctgt cagtgagatg gcttctaaca agttcaaaag aatgctgaac
 721 cgggagctga cacacctctc agagatgagc cgatcaggga accaggtgtc tgaatacatt
 781 tcaaatactt tcttagacaa gcagaatgat gtggagatcc catctcctac ccagaaagac
 841 agggagaaaa agaaaaagca gcagctcatg acccagataa gtggagtgaa gaaattaatg
 901 catagttcaa gcctaaacaa tacaagcatc tcacgctttg gagtcaacac tgaaaatgaa
 961 gatcacctgg ccaaggagct ggaagacctg aacaaatggg gtcttaacat ctttaatgtg
1021 gctggatatt ctcacaatag acccctaaca tgcatcatgt atgctatatt ccaggaaaga
1081 gacctcctaa agacattcag aatctcatct gacacattta taacctacat gatgacttta
1141 gaagaccatt accattctga cgtggcatat cacaacagcc tgcacgctgc tgatgtagcc
1201 cagtcgaccc atgttctcct ttctacacca gcattagacg ctgtcttcac agatttggag
1261 atcctggctg ccatttttgc agctgccatc catgacgttg atcatcctgg agtctccaat
1321 cagtttctca tcaacacaaa ttcagaactt gctttgatgt ataatgatga atctgtgttg
1381 gaaaatcatc accttgctgt gggtttcaaa ctgctgcaag aagaacactg tgacatcttc
1441 atgaatctca ccaagaagca gcgtcagaca ctcaggaaga tggttattga catggtgtta
1501 gcaactgata tgtctaaaca tatgagcctg ctggcagacc tgaagacaat ggtagaaacg
1561 aagaaagtta caagttcagg cgttcttctc ctagacaact ataccgatcg cattcaggtc
1621 cttcgcaaca tggtacactg tgcagacctg agcaacccca ccaagtcctt ggaattgtat
1681 cggcaatgga cagaccgcat catggaggaa tttttccagc agggagacaa agagcgggag
1741 aggggaatgg aaattagccc aatgtgtgat aaacacacag cttctgtgga aaaatcccag
1801 gttggtttca tcgactacat tgtccatcca ttgtgggaga catgggcaga tttggtacag
1861 cctgatgctc aggacattct cgatacctta gaagataaca ggaactggta tcagagcatg
1921 atacctcaaa gtccctcacc accactggac gagcagaaca gggactgcca gggtctgatg
1981 gagaagtttc agtttgaact gactctcgat gaggaagatt ctgaaggacc tgagaaggag
2041 ggagagggac acagctattt cagcagcaca aagacgcttt gtgtgattga tccagaaaac
2101 agagattccc tgggagagac tgacatagac attgcaacag aagacaagtc ccccgtggat
2161 aca
```

Figure 28

PDE4B3 Protein sequence

EESYQKLAMETLEELDWCLDQLETIQTYRSVSEMASNKFK
RMLNRELTHLSEMSRSGNQVSEYISNTFLDKQNDVEIPSPTQKDREKKKKQQLMTQISG
VKKLMHSSSLNNTSISRFGVNTENEDHLAKELEDLNKWGLNIFNVAGYSHNRPLTCIMY
AIFQERDLLKTFRISSDTFITYMMTLEDHYHSDVAYHNSLHAADVAQSTHVLLSTPALD
AVFTDLEILAAIFAAAIHDVDHPGVSNQFLINTNSELALMYNDESVLENHHLAVGFKLL
QEEHCDIFMNLTKKQRQTLRKMVIDMVLATDMSKHMSLLADLKTMVETKKVTSSGVLLL
DNYTDRIQVLRNMVHCADLSNPTKSLELYRQWTDRIMEEFFQQGDKERERGMEISPMCD
KHTASVEKSQVGFIDYIVHPLWETWADLVQPDAQDILDTLEDNRNWYQSMIPQSPSPPL
DEQNRDCQGLMEKFQFELTLDEEDSEGPEKEGEGHSYFSSTKTLCVIDPENRDSLGETD
IDIATEDKSPVDT

Figure 29

PDE4B2 (acc. NM_002600) Nucleic acid seqeunce

```
   1 gaattcctcc tctcttcacc ccgttagctg ttttcaatgt aatgctgccg tccttctctt
  61 gcactgcctt ctgcgctaac acctccattc ctgtttataa ccgtgtattt attacttaat
 121 gtatataatg taatgttttg taagttatta atttatatat ctaacattgc ctgccaatgg
 181 tggtgttaaa tttgtgtaga aaactctgcc taagagttac gacttttttct tgtaatgttt
 241 tgtattgtgt attatataac ccaaacgtca cttagtagag acatatggcc cccttggcag
 301 agaggacagg ggtgggcttt tgttcaaagg gtctgccctt tccctgcctg agttgctact
 361 tctgcacaac ccctttatga accagttttc acccgaattt tgactgtttc atttagaaga
 421 aaagcaaaat gagaaaaagc tttcctcatt tctccttgag atggcaaagc actcagaaat
 481 gacatcacat accctaaaga accctgggat gactaaggca gagagagtct gagaaaactc
 541 tttggtgctt ctgcctttag ttttaggaca catttatgca gatgagctta taagagaccg
 601 ttcccctccgc cttcttcctc agaggaagtt tcttggtaga tcaccgacac ctcatccagg
 661 cgggggggttg gggggaaact tggcaccagc catcccaggc agagcaccac tgtgatttgt
 721 tctcctggtg gagagagctg gaaggaagga gccagcgtgc aaataatgaa ggagcacggg
 781 ggcaccttca gtagcaccgg aatcagcggt ggtagcggtg actctgctat ggacagcctg
 841 cagccgctcc agcctaacta catgcctgtg tgtttgtttg cagaagaatc ttatcaaaaa
 901 ttagcaatgg aaacgctgga ggaattagac tggtgtttag accagctaga gaccatacag
 961 acctaccggt ctgtcagtga gatggcttct aacaagttca aaagaatgct gaaccgggag
1021 ctgacacacc tctcagagat gagccgatca gggaaccagg tgtctgaata catttcaaat
1081 actttcttag acaagcagaa tgatgtggag atcccatctc ctacccagaa agacagggag
1141 aaaaagaaaa agcagcagct catgacccag ataagtggag tgaagaaatt aatgcatagt
1201 tcaagcctaa acaatacaag catctcacgc tttggagtca acactgaaaa tgaagatcac
1261 ctggccaagg agctggaaga cctgaacaaa tggggtctta acatctttaa tgtggctgga
1321 tattctcaca atagacccct aacatgcatc atgtatgcta tattccagga aagagacctc
1381 ctaaagacat tcagaatctc atctgacaca tttataacct acatgatgac tttagaagac
1441 cattaccatt ctgacgtggc atatcacaac agcctgcacg ctgctgatgt agcccagtcg
1501 acccatgttc tcctttctac accagcatta gacgctgtct tcacagattt ggagatcctg
1561 gctgccattt ttgcagctgc catccatgac gttgatcatc ctggagtctc caatcagttt
1621 ctcatcaaca caaattcaga acttgctttg atgtataatg atgaatctgt gttggaaaat
1681 catcaccttg ctgtgggttt caaactgctg caagaagaac actgtgacat cttcatgaat
1741 ctcaccaaga agcagcgtca gacactcagg aagatggtta ttgacatggt gttagcaact
1801 gatatgtcta aacatatgag cctgctggca gacctgaaga caatggtaga aacgaagaaa
1861 gttacaagtt caggcgttct tctcctagac aactataccg atcgcattca ggtccttcgc
1921 aacatggtac actgtgcaga cctgagcaac cccaccaagt ccttggaatt gtatcggcaa
1981 tggacagacc gcatcatgga ggaatttttc cagcagggag acaaagagcg ggagagggga
2041 atggaaatta gcccaatgtg tgataaacac acagcttctg tggaaaaatc ccaggttggt
2101 ttcatcgact acattgtcca tccattgtgg gagacatggg cagatttggt acagcctgat
2161 gctcaggaca ttctcgatac cttagaagat aacaggaact ggtatcagag catgatacct
2221 caaagtccct caccaccact ggacgagcag aacagggact gccagggtct gatggagaag
2281 tttcagtttg aactgactct cgatgaggaa gattctgaag gacctgagaa ggagggagag
2341 ggacacagct atttcagcag cacaaagacg ctttgtgtga ttgatccaga aaacagagat
2401 tccctgggag agactgacat agacattgca acagaagaca agtcccccgt ggatacataa
2461 tccccctctc cctgtggaga tgaacattct atccttgatg agcatgccag ctatgtggta
2521 gggccagccc accatggggg ccaagacctg cacaggacaa gggccacctg gcctttcagt
2581 tacttgagtt tggagtcaga aagcaagacc aggaagcaaa tagcagctca ggaaatccca
2641 cggttgactt gccttgatgg caagcttggt ggagagggct gaagctgttg ctgggggccg
2701 attctgatca agacacatgg cttgaaaatg gaagacacaa aactgagaga tcattctgca
2761 ctaagtttcg ggaacttatc cccgacagtg actgaactca ctgactaata acttcattta
2821 tgaatctttct cacttgtccc tttgtctgcc aacctgtgtg ccttttttgt aaaacatttt
2881 catgtcttta aaatgcctgt tgaatacctg gagtttagta tcaacttcta cacagataag
2941 ctttcaaagt tgacaaactt ttttgactct ttctggaaaa gggaaagaaa atagtcttcc
```

```
3001 ttctttcttg ggcaatatcc ttcactttac tacagttact tttgcaaaca gacagaaagg
3061 atacacttct aaccacattt tacttccttc ccctgttgtc cagtccaact ccacagtcac
3121 tcttaaaact tctctctgtt tgcctgcctc caacagtact tttaactttt tgctgtaaac
3181 agaataaaat tgaacaaatt agggggtaga aaggagcagt ggtgtcgttc accgtgagag
3241 tctgcataga actcagcagt gtgccctgct gtgtcttgga ccctgccccc cacaggagtt
3301 gctacagtcc ctggccctgc ttcccatcct cctctcttca ccccgttagc tgttttcaat
3361 gtaatgctgc cgtccttctc ttgcactgcc ttctgcgcta acacctccat tcctgtttat
3421 aaccgtgtat ttattactta atgtatataa tgtaatgttt tgtaagttat taatttatat
3481 atctaacatt gcctgccaat ggtggtgtta aatttgtgta gaaaactctg cctaagagtt
3541 acgacttttt cttgtaatgt tttgtattgt gtattatata acccaaacgt cacttagtag
3601 agacatatgg ccccccttggc agagaggaca ggggtgggct tttgttcaaa gggtctgccc
3661 tttccctgcc tgagttgcta cttctgcaca acccctttat gaaccagttt tggaaacaat
3721 attctcacat tagatactaa atggtttata ctgagtcttt tacttttgta tagcttgata
3781 ggggcagggg caatgggatg tagtttttac ccaggttcta tccaaatcta tgtgggcatg
3841 agttgggtta taactggatc ctactatcat tgtggctttg gttcaaaagg aaacactaca
3901 tttgctcaca gatgattctt ctgattcttc tgaatgctcc cgaactactg actttgaaga
3961 ggtagcctcc tgcctgccat taagcaggaa tgtcatgttc cagttcatta caaaagaaaa
4021 caataaaaca atgtgaattt ttataataaa aaaaaaaaaa aggaattc
```

Figure 30

PDE4B2 Protein sequence

```
MKEHCGTPSSTGESGGSGDSAMDSLQPLQPNYMPVCLFAEESYQKLAMETLEELDWCLD
QLETIQTYRSVSEMASNKFKRMLNRELTHLSEMSRSGNQVSEYISNTFLDKQNDVEIPS
PTQKDREKKKKQQLMTQISGVKKLMHSSSLNNTSISRFGVNTENEDHLAKELEDLNKWG
LNIFNVAGYSHNRPLTCIMYAIFQERDLLKTFRISSDTFITYMMTLEDHYHSDVAYHNS
LHAADVAQSTHVLLSTPALDAVFTDLEILAAIFAAAIHDVDHPGVSNQFLINTNSELAL
MYNDESVLENHHLAVGFKLLQEEHCDIFMNLTKKQRQTLRKMVIDMVLATDMSKHMSLL
ADLKTMVETKKVTSSGVLLLDNYTDRIQVLRNMVHCADLSNPTKSLELYRQWTDRIMEE
FFQQGDKERERGMEISPMCDKHTASVEKSQVGFIDYIVHPLWETWADLVQPDAQDILDT
LEDNRNWYQSMIPQSPSPPLDEQNRDCQGLMEKFQFELTLDEEDSEGPEKEGEGHSYFS
STKTLCVIDPENRDSLGETDIDIATEDKSPVDT
```

Figure 31

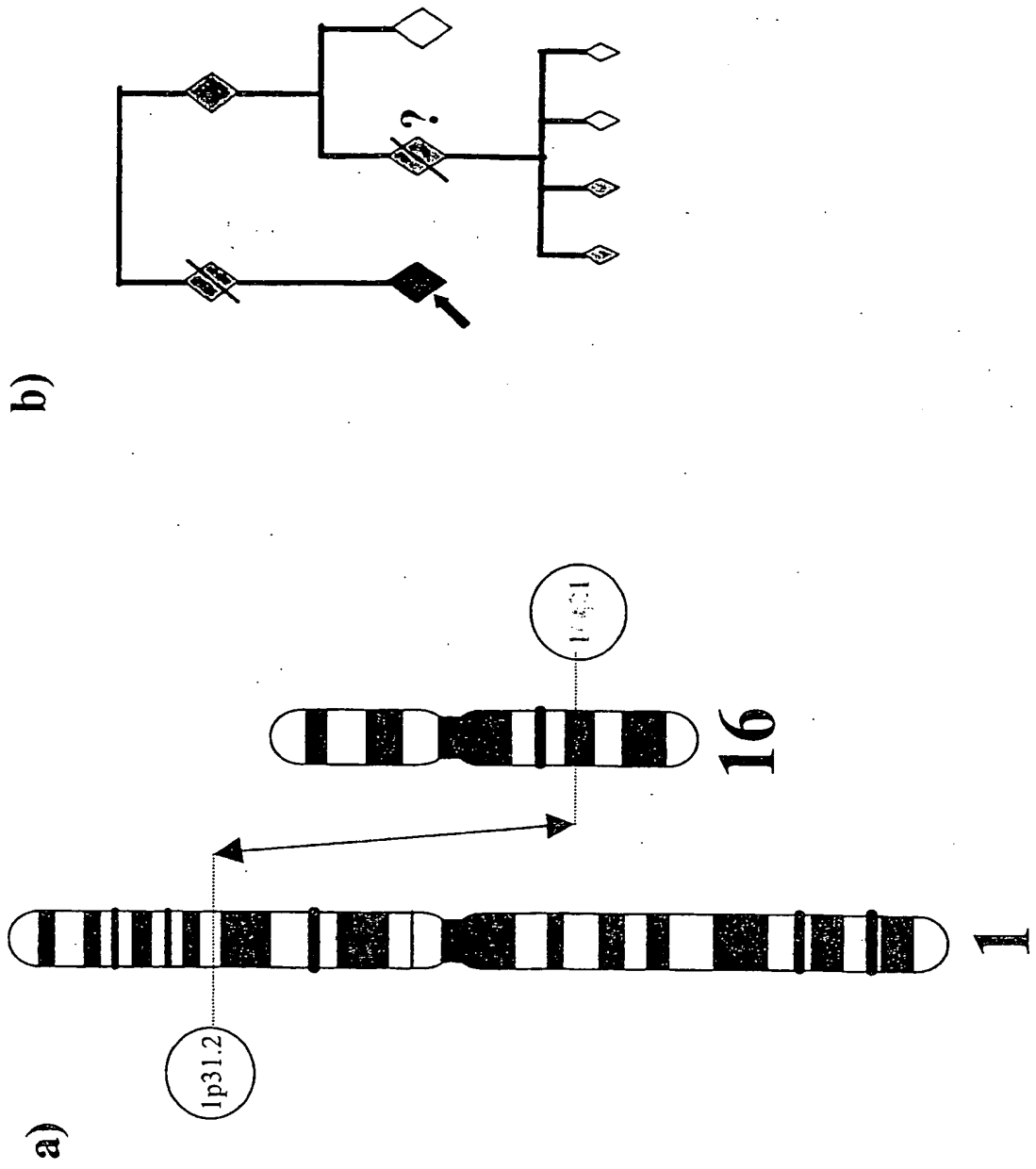


Figure 32

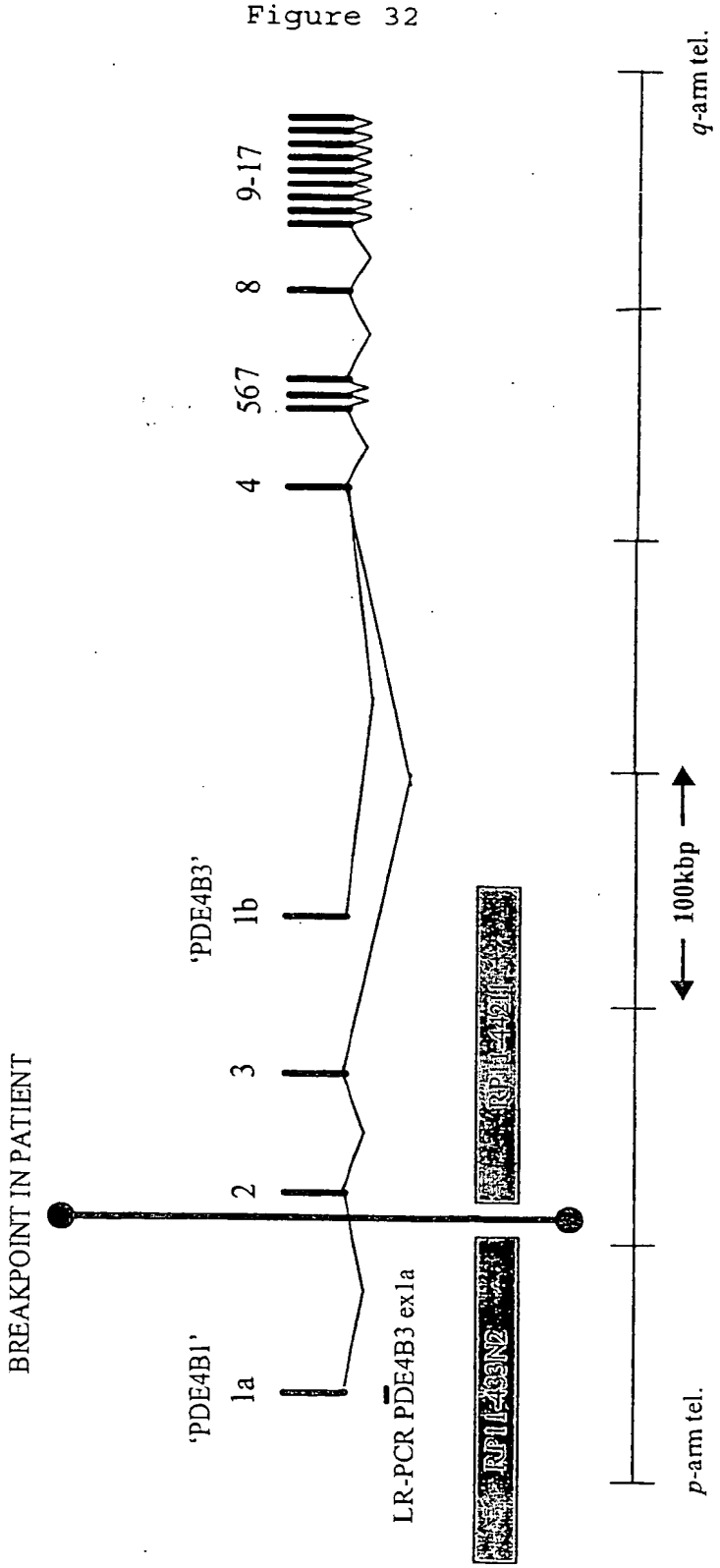

BREAKPOINT IN PATIENT
'PDE4B1'
1a
2
3
'PDE4B3'
1b
4
567
8
9-17
LR-PCR PDE4B3 ex1a
RP11-421F4
RP11-243N2
100kbp
q-arm tel.
p-arm tel.

Figure 33

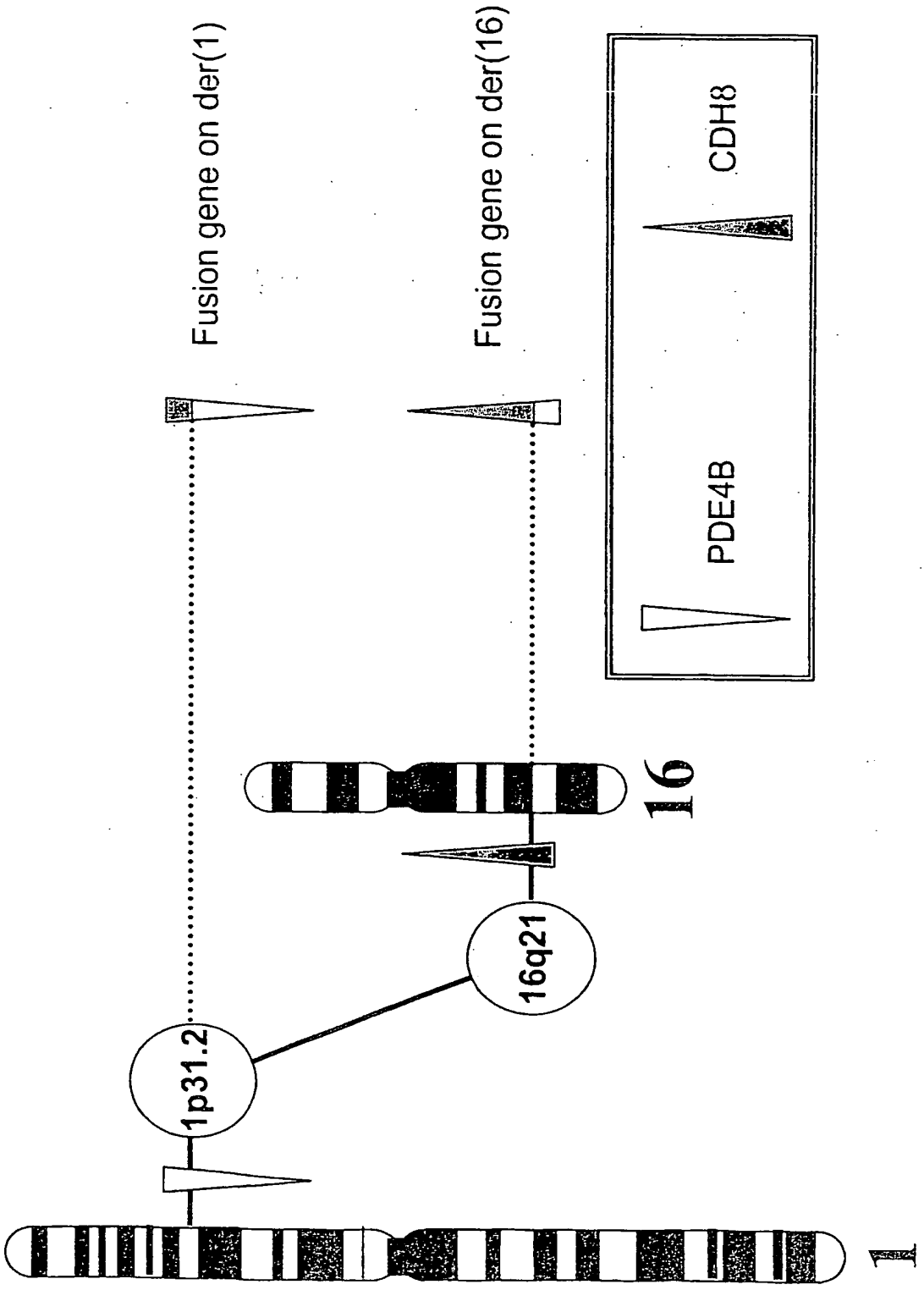

Fusion gene on der(1)
Fusion gene on der(16)
CDH8
PDE4B
16
16q21
1p31.2
1

Figure 34

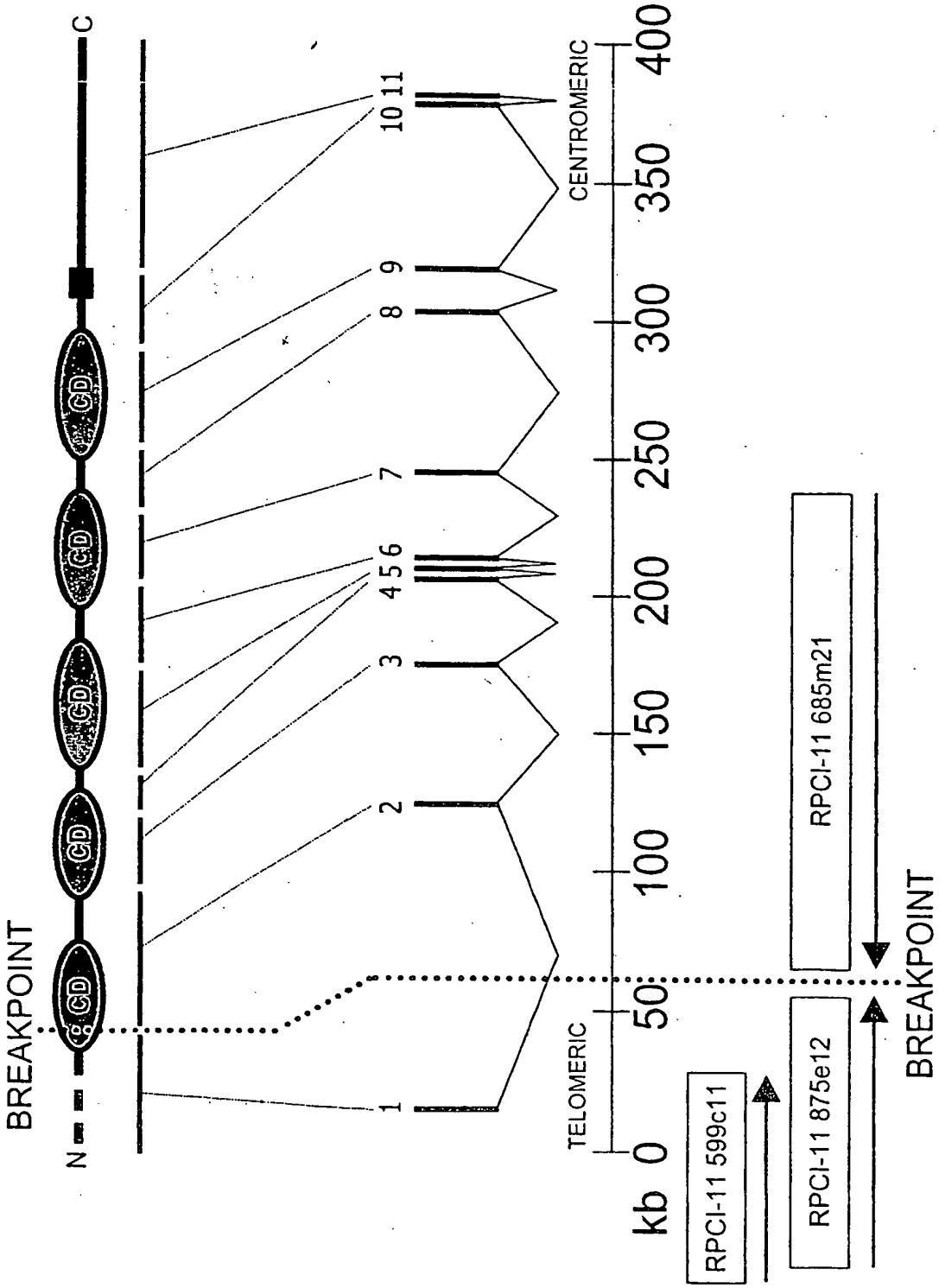

C
N
CD
CD
CD
CD
CD
BREAKPOINT
1
2
3
4 5 6
7
8
9
10 11
CENTROMERIC
TELOMERIC
kb 0
50
100
150
200
250
300
350
400
RPCI-11 599c11
RPCI-11 875e12
RPCI-11 685m21
BREAKPOINT

Figure 35

```
   1 agccatttgt gaacctggag gcttgacatt cgccagcgca gggccccaca agagaaattt
  61 caatgaaaag aaaagccaat ggattgtggt cttagaaaag ctgcttagat gatgtctgtt
 121 tcccgtgcta tagacacgtg gcagagctgt aagtaaatgc tcggcactgc atgatgaatt
 181 ggatggctgc agaccggaga caaaaaaaat aattgtctca ttttcgtggt gatttgctta
 241 actggtggga ccatgccaga acggctagcg gaaatgctct tggatctctg gactccatta
 301 ataatattat ggattactct tccccttgc atttacatgg ctccgatgaa tcagtctcaa
 361 gttttaatga gtggatcccc tttggaacta aacagtctgg gtgaagaaca gcgaattttg
 421 aaccgctcca aaagaggctg ggtttggaat caaatgtttg tcctggaaga gttttctgga
 481 cctgaaccga ttcttgttgg ccggctacac acagacctgg atcctgggag caaaaaaatc
 541 aagtatatcc tatcaggtga tggagctggg accatatttc aaataaatga tgtaactgga
 601 gatatccatg ctataaaaag acttgaccgg gaggaaaagg ctgagtatac cctaacagct
 661 caagcagtgg actgggagac aagcaaacct ctggagcctc cttctgaatt tattattaaa
 721 gttcaagaca tcaatgacaa tgcaccagag tttcttaatg gaccctatca tgctactgtg
 781 ccagaaatgt ccattttggg tacatctgtc actaacgtca ctgcgaccga cgctgatgac
 841 ccagtttatg gaaacagtgc aaagttggtt tatagtatat tggaagggca gccttatttt
 901 tccattgagc ctgaaacagc tattataaaa actgcccttc ccaacatgga cagagaagcc
 961 aaggaggagt acctggttgt tatccaagcc aaagatatgg gtggacactc tggtggcctg
1021 tctgggacca cgacacttac agtgactctt actgatgtta atgacaatcc tccaaaattt
1081 gcacagagcc tgtatcactt ctcagtaccg gaagatgtgg ttcttggcac tgcaatagga
1141 agggtgaagg ccaatgatca ggatattggt gaaaatgcac agtcatcata tgatatcatc
1201 gatggagatg gaacagcact ttttgaaatc acttctgatg cccaggccca ggatggcatt
1261 ataaggctaa gaaaacctct ggactttgag accaaaaaat cctatacgct aaaggtagag
1321 gcagccaatg tccatattga cccacgcttc agtggcaggg ggccctttaa agacacggcg
1381 acagtcaaaa tcgtggttga agatgctgat gagcctccgg tcttctcttc accgacttac
1441 ctacttgaag ttcatgaaaa tgctgctcta aactccgtga ttgggcaagt gactgctcgt
1501 gaccctgata tcacttccag tcctataagg ttttccatcg accggcacac tgacctggag
1561 aggcagttca acattaatgc agacgatggg aagataacgc tggcaacacc acttgacaga
1621 gaattaagtg tatggcacaa cataacaatc attgctactg aaattaggaa ccacagtcag
1681 atatcacgag tacctgttgc tattaaagtg ctggatgtca atgacaacgc ccctgaattc
1741 gcatccgaat atgaggcatt tttatgtgaa aatggaaaac ccggccaagt cattcaaact
1801 gttagcgcca tggacaaaga tgatcccaaa aacggacatt atttcttata cagtctcctt
1861 ccagaaatgg tcaacaatcc gaatttcacc atcaagaaaa atgaagataa ttccctcagt
1921 attttggcaa agcataatgg attcaaccgc cagaagcaag aagtctatct tttaccaatc
1981 ataatcagtg atagtggaaa tcctccactg agcagcacta gcaccttgac aatcagggtc
2041 tgtggctgca gcaatgacgg tgtcgtccag tcttgcaatg tcgaagctta tgtccttcca
2101 attggactca gtatgggcgc cttaattgcc atattagcat gcatcatttt gctgttagtc
2161 atcgtggtgc tgtttgtaac tctacggcgg cataaaaatg aaccattaat tatcaaagat
2221 gatgaagacg ttcgagaaaa catcattcgc tacgatgatg aaggaggagg ggaggaggac
2281 acagaggctt ttgacattgc aactttacaa aatccagatg gaattaatgg atttttaccc
2341 cgtaaggata ttaaaccaga tttgcagttt atgccaaggc aagggcttgc tccagttcca
2401 aatggtgttg atgtcgatga atttataaat gtaaggctgc atgaggcaga taatgatccc
2461 acggccccgc catatgactc cattcagata tatggctatg aaggccgagg gtcagtggct
2521 ggctccctca gctccttgga gtccaccaca tcagactcag accagaattt tgactacctc
2581 agtgactggg gtccccgctt taagagactg ggcgaactct actctgttgg tgaaagtgac
2641 aaagaaactt gacagtggat tataaataaa tcactggaac tgagcattct gtaatattct
2701 agggtcactc cccttagata caaccaatgt ggctatttgt tttagaggca agtttagcac
2761 cagtcatcta taaactcaac cacattttaa tgttgaacca aaaaaagata ataaaataaa
2821 aaagtatatg ttaggaggtt ataaatcttg tggagtgtga attaagtatg tggagtgtct
2881 agaagtcctt ggatatttga tatttacctg accaccacag acaaagatt
```

Figure 36

1 *MPERLAEMLL DLWTPLIILW ITLPPCIYMA* PMNQSQVLMS GSPLELNSLG EEQRILNRSK
61 RGWVWNQMFV LEEFSGPEPI LVGRLHTDLD PGSKKIKYIL SGDGAGTIFQ INDVTGDIHA
121 IKRLDREEKA EYTLTAQAVD WETSKPLEPP SEFIIKVQDI NDNAPEFLNG PYHATVPEMS
181 ILGTSVTNVT ATDADDPVYG NSAKLVYSIL EGQPYFSIEP ETAIIKTALP NMDREAKEEY
241 LVVIQAKDMG GHSGGLSGTT TLTVTLTDVN DNPPKFAQSL YHFSVPEDVV LGTAIGRVKA
301 NDQDIGENAQ SSYDIIDGDG TALFEITSDA QAQDGIIRLR KPLDFETKKS YTLKVEAANV
361 HIDPRFSGRG PFKDTATVKI VVEDADEPPV FSSPTYLLEV HENAALNSVI GQVTARDPDI
421 TSSPIRFSID RHTDLERQFN INADDGKITL ATPLDRELSV WHNITIIATE IRNHSQISRV
481 PVAIKVLDVN DNAPEFASEY EAFLCENGKP GQMIQTVSAM DKDDPKNGHY FLYSLLPEMV
541 NNPNFTIKKN EDNSLSILAK HNGFNRQKQE VYLLPIIISD SGNPPLSSTS TLTIRVCGCS
601 NDGVVQSCNV EAYVLPIGLS MGALIAILAC IILLLVIVVL FVTLRRHKNE PLIIKDDEDV
661 RENIIRYDDE GGGEEDTEAF DIATLQNPDG INGFLPRKDI KPDLQFMPRQ GLAPVPNGVD
721 VDEFINVRLH EADNDPTAPP YDSIQIYGYE GRGSVAGSLS SLESTTSDSD QNFDYLSDWG
781 PRFKRLGELY SVGESDKET

Figure 37

a)

MPERLAEMLLDLWTPLIILWITLPPCIYMAPMNQSQVLMSGSPLELNSLGEEQRILNRS

KRGWVWNQMFVLEEFSGPEPILVGRVLKSVSKLH*

b)

```
      G  R  G  G  A  A  E  A  P  R  A  G  G  G  R  L  L  R  G  Q
3     ggccgcggcggtgcagcagaggcgcctcgggcaggaggagggcggcttctgcgagggcag  62
      P  E  L  H  T  D  L  D  P  G  S  K  K  I  K  Y  I  L  S  G
63    cctgagctacacacagacctggatcctgggagcaaaaaaatcaagtatatcctatcaggt  122
      D  G  A  G  T  I  F  Q  I  N  D  V  T  G  D  I  H  A  I  K
123   gatggagctgggaccatatttcaaataaatgatgtaactggagatatccatgctataaaa  182
      R  L  D  R  E  E  K  A  E  Y  T  L  T  A  Q  A  V  D  W  E
183   agacttgaccgggaggaaaaggctgagtataccctaacagctcaagcagtggactgggag  242
      T  S  K  P  L  E  P  P  S  E  F  I  I  K  V  Q  D  I  N  D
243   acaagcaaacctctggagcctccttctgaatttattattaaagttcaagacatcaatgac  302
      N  A  P  E  F  L  N  G  P  Y  H  A  T  V  P  E  M  S  I  L
303   aatgcaccagagtttcttaatggaccctatcatgctactgtgccagaaatgtccattttg  362
      G  T  S  V  T  N  V  T  A  T  D  A  D  D  P  V  Y  G  N  S
363   ggtacatctgtcactaacgtcactgcgaccgacgctgatgacccagtttatggaaacagt  422
      A  K  L  V  Y  S  I  L  E  G  Q  P  Y  F  S  I  E  P  E  T
423   gcaaagttggtttatagtatattggaagggcagccttatttttccattgagcctgaaaca  482
      A  I  I  K  T  A  L  P  N  M  D  R  E  A  K  E  E  Y  L  V
483   gctattataaaaactgcccttccaacatggacagagaagccaaggaggagtacctggtt  542
      V  I  Q  A  K  D  M  G  G  H  S  G  G  L  S  G  T  T  T  L
543   gttatccaagccaaagatatgggtggacactctggtggcctgtctgggaccacgacactt  602
      T  V  T  L  T  D  V  N  D  N  P  P  K  F  A  Q  S  L  Y  H
603   acagtgactcttactgatgttaatgacaatcctccaaaatttgcacagagcctgtatcac  662
      F  S  V  P  E  D  V  V  L  G  T  A  I  G  R  V  K  A  N  D
663   ttctcagtaccggaagatgtggttcttggcactgcaataggaagggtgaaggccaatgat  722
      Q  D  I  G  E  N  A  Q  S  S  Y  D  I  I  D  G  D  G  T  A
723   caggatattggtgaaaatgcacagtcatcatatgatatcatcgatggagatggaacagca  782
      L  F  E  I  T  S  D  A  Q  A  Q  D  G  I  I  R  L  R  K  P
783   ctttttgaaatcacttctgatgcccaggcccaggatggcattataaggctaagaaaacct  842
      L  D  F  E  T  K  K  S  Y  T  L  K  V  E  A  A  N  V  H  I
843   ctggactttgagaccaaaaaatcctatacgctaaaggtagaggcagccaatgtccatatt  902
      D  P  R  F  S  G  R  G  P  F  K  D  T  A  T  V  K  I  V  V
903   gacccacgcttcagtggcaggggggccctttaaagacacggcgacagtcaaaatcgtggtt  962
      E  D  A  D  E  P  P  V  F  S  S  P  T  Y  L  L  E  V  H  E
963   gaagatgctgatgagcctccggtcttctcttcaccgacttacctacttgaagttcatgaa 1022
      N  A  A  L  N  S  V  I  G  Q  V  T  A  R
1023  aatgctgctctaaactccgtgattgggcaagtgactgctcgt etc.....
```